

91473

STIC-Biotech/ChemLib

From: Slobodyansky, Elizabeth
Sent: Monday, April 14, 2003 11:43 AM
To: STIC-Biotech/ChemLib
Subject: 09/527,376

Please search for case 09/527,376:



SEQ ID NOs: 1 and 2 against commercial and interference databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner

Art Unit 1652
CM1 10D11
703-306-3222

mail box 10D01.

Mary Jane Ruhl
Tech. Info. Specialist, STIC
TC-1600
CM-1, Room 6A-06
Phone: 605-1155

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 4/22/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 18:32:39 ; Search time 22 Seconds
(without alignments)
251.432 Million cell updates/sec

Title: us-09-527-376-2

Perfect score: 981

Sequence: 1 MTAPSCAFFPQFRQPSVSL.....SPVGMPIYKEVRLMIPL 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	511	52.1	198	4	US-09-045-973-5
2	275	28.0	226	4	US-09-045-973-8
3	252	25.7	367	2	US-08-530-290-24
4	252	25.7	367	2	US-08-990-379-6
5	240	24.5	394	2	US-08-530-290-23
6	236	24.1	393	2	US-08-990-379-4
7	236	24.1	395	2	US-08-990-379-5
8	222	22.6	176	4	US-09-704-139-2
9	221	22.5	397	2	US-08-990-379-8
10	220	22.4	173	4	US-09-704-139-4
11	216	22.0	117	5	US-07-988-273-4
12	216	22.0	117	5	PCT-US93-12019-4
13	213.5	21.8	172	4	US-09-704-139-5
14	208	21.2	313	2	US-08-990-379-7
15	208	21.2	314	4	US-09-164-193-22
16	208	21.2	314	4	US-09-221-448A-22
17	194.5	19.8	198	4	US-09-163-833-2
18	179.5	18.3	207	4	US-09-013-881-4
19	171	17.4	263	4	US-09-164-193-5
20	171	17.4	263	4	US-09-221-448A-5
21	162.5	16.6	185	1	US-07-988-273-2
22	162.5	16.6	185	3	US-08-848-810-25
23	162.5	16.6	185	4	US-09-164-193-21
24	162.5	16.6	185	4	US-09-221-448A-21
25	162.5	16.6	185	5	PCT-US93-12019-2
26	146	14.9	72	2	US-08-530-290-19
27	141	14.4	353	4	US-09-013-881-3

Sequence 20, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 14, Appl
Sequence 21, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 16, Appl
Sequence 15, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 2, Appl
Sequence 5, Appl
Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-045-973-5
Sequence 5, Application US/09045973
Patent No. 6165767
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTN016
CLONE: 3041794
US-09-045-973-5


```
Db 163 SCSTPLY----DOGGVEILPFLYLGSAHASKDMLDALGITALINVSANCPNHFEGHY 218
QY 65 QYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSALCLAYLMKYHA 124
Db 219 QYKSIPEVDNHHKADISSWNEAIDFDSIKNAGGRVHVHCQAGISRSATICLAYLMRTNR 278
QY 125 MSLLDAHTWTKSCRPPIRPNNSGFEQLIHFEFOLFGKNTVHMVSSPVGMIPD 176
Db 279 VKLDEAFEFVKQRSSIISPNFSPMGQLLOFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 4
US-08-990-379-6
; Sequence 6, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Phillip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; EARLIER FILING DATE: 1997-12-15
; EARLIER FILING DATE: 1996-06-14
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-990-379-6

Query Match 25.7%; Score 252; DB 2; Length 367;
Best Local Similarity 34.3%; Pred. No. 1.4e-22;
Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

QY 5 SCAPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITPMVINSVVNTLYEDI 64
Db 163 SCSTPLY----DOGGVEILPFLYLGSAHASKDMLDALGITALINVSANCPNHFEGHY 218
QY 65 QYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSALCLAYLMKYHA 124
Db 219 QYKSIPEVDNHHKADISSWNEAIDFDSIKDAGGRVHVHCQAGISRSATICLAYLMRTNR 278
QY 125 MSLLDAHTWTKSCRPPIRPNNSGFEQLIHFEFOLFGKNTVHMVSSPVGMIPD 176
Db 279 VKLDEAFEFVKQRSSIISPNFSPMGQLLOFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 5
US-08-530-290-23
; Sequence 23, Application US/08530290
; Patent No. 5938721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-000000005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-530-290-23

Query Match 24.5%; Score 240; DB 2; Length 394;
Best Local Similarity 33.9%; Pred. No. 4.8e-21;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

QY 5 SCAPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITPMVINSVVNTLYEDI 64
Db 185 SCTPLHDQE---GPVEILPFLYLGSAHARRDMLDALGITALLNVSDCPNHFEGHY 240
QY 65 QYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSALCLAYLMKYHA 124
Db 241 QYKCIPEVDNHHKADISSWNEAIEYIDAVKDCRGRVHVHCQAGISRSATICLAYLMKKR 300
QY 125 MSLLDAHTWTKSCRPPIRPNNSGFEQLIHFEFOLFGKNTVHMVSSPVG 172
Db 301 VRLEAEFEVKQRSSIISPNFSPMGQLLOFESQVLATSCAAEAASPSG 348

RESULT 6
US-08-990-379-4
; Sequence 4, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Phillip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-990-379-4

Query Match 24.1%; Score 236; DB 2; Length 393;
```


FILE REFERENCE: 10448-018001
CURRENT APPLICATION NUMBER: US/09/704,139
CURRENT FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: US 60/185,772
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 173
TYPE: PRT
ORGANISM: Artificial/Unknown
FEATURE: VARIANT
NAME/KEY: (1)..(173)
LOCATION: (1)..(173)
OTHER INFORMATION: consensus sequence
US-09-704-139-4

Query Match 22.4%; Score 220; DB 4; Length 173;
Best Local Similarity 34.9%; Pred. No. 4e-19;
Matches 60; Conservative 21; Mismatches 55; Indels 36; Gaps 5;
QY 19 GLSQTSLYISNGVAAN--NKLMLSSNOITMVINVSVEVNTLYED----- 63
Db 1 GPSEILPHLYLGSYSTASEANLALKLKGITHVINTEVPNPFELDKKNDRHYNAYIS 60
QY 64 ----IQMQVP-----VADSPNSRLCDFDPIADHIHVSVMKQGRITLLHCAAGVSR 110
Db 61 KNSGFTYLIQPNVDHIIYHIAWNETKISKYFDEAVDFIDARQKGGKVLVHCQAGISR 120
QY 111 SAAALCLAYLMKYHMSLLDA-----HPWTKSCRPIIRPNSGFWEQLIHYE 155
Db 121 SATLIAYLMKTRNLNSLAYDFVYVYHIKERRC-PIISPNGFLRLQLEYE 171

RESULT 11
US-07-988-273-4
Sequence 4, Application US/07988273
Patent No. 5512434
GENERAL INFORMATION:
APPLICANT: AARONSON, Stuart A.
APPLICANT: BOTTARO, Donald P.
APPLICANT: ISHIBASHI, Toshio
APPLICANT: MIKI, Toru
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,273
FILING DATE: 19921214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/182 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: AMINO ACID

TOPOLOGY: linear
US-07-988-273-4
Query Match 22.0%; Score 216; DB 1; Length 117;
Best Local Similarity 39.7%; Pred. No. 6.8e-19;
Matches 46; Conservative 20; Mismatches 50; Indels 0; Gaps 0;
QY 40 MLSSNOITMVINVSVEVNTLYEDIQMQVPVADSPNSRLCDFDPIADHIHVSVMKQGR 99
Db 1 MLDALGITALINVSANCPNHFEGHYQKIPVEDNHNKADISSWFNEAIDFIDSIKDAGGR 60
QY 100 TLLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYE 155
Db 61 VFVHCQAGISRATICLAYLMRTNRVKLDEAFEVVKQRRSIIISPNFSFMGQLLOQE 116
RESULT 12
PCT-US93-12019-4
Sequence 4, Application PC/TUS9312019
GENERAL INFORMATION:
APPLICANT: EXPRESSION CLONING OF A HUMAN
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12019
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,273
FILING DATE: 14-DEC-1992
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US93-12019-4
Query Match 22.0%; Score 216; DB 5; Length 117;
Best Local Similarity 39.7%; Pred. No. 6.8e-19;
Matches 46; Conservative 20; Mismatches 50; Indels 0; Gaps 0;
QY 40 MLSSNOITMVINVSVEVNTLYEDIQMQVPVADSPNSRLCDFDPIADHIHVSVMKQGR 99
Db 1 MLDALGITALINVSANCPNHFEGHYQKIPVEDNHNKADISSWFNEAIDFIDSIKDAGGR 60
QY 100 TLLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYE 155
Db 61 VFVHCQAGISRATICLAYLMRTNRVKLDEAFEVVKQRRSIIISPNFSFMGQLLOQE 116
RESULT 13
US-09-704-139-5
Sequence 5, Application US/09704139
Patent No. 6420153
GENERAL INFORMATION:
APPLICANT: Kapeller, Rosanna
TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREOF
FILE REFERENCE: 10448-018001
CURRENT APPLICATION NUMBER: US/09/704,139
CURRENT FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: US 60/185,772
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 172
TYPE: PRT
ORGANISM: Artificial/Unknown

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(172)
; OTHER INFORMATION: consensus sequence
US-09-704-139-5

Query Match 21.8%; Score 213.5; DB 4; Length 172;
Best Local Similarity 34.5%; Pred. No. 2.5e-18;
Matches 59; Conservative 21; Mismatches 56; Indels 35; Gaps 5;
Qy 19 GLSQITKSLYISNGVAAN--NKLMLSSNQITWVNVSEVVNTLY----- 61
Db 1 GPSELPHLYGSYSDASEANLALKLIGTHVINVTVEENFELKKKNDRYTNEYIS 60
Qy 62 --EDIQMVDP-----VADSPNSRLCDFDPIADHIHSVEMKQGRTHLHCAAGVRS 111
Db 61 KSGFTYLIQIPNVDDIYHIAWNTETKISKYLEEAEVIEDAEKKGKVLVHCOAGVRS 120
Qy 112 AALCLAYLMKYHAMSLDLA-----HTWTKSCRPIIRPNSGFWQLIHYE 155
Db 121 ATLVIAYLMKTRNLSLDRDAYDEVVYVYHIKERRC--PIISPNFGFLRLIEYE 170

RESULT 14

US-08-990-379-7
; Sequence 7, Application US/08990379
; Patent No. 5998188

GENERAL INFORMATION:

; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; TITLE OF INVENTION: Their Biologically Active Expression Products
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-990-379-7

Query Match 21.2%; Score 208; DB 2; Length 313;
Best Local Similarity 33.1%; Pred. No. 3e-17;
Matches 55; Conservative 26; Mismatches 75; Indels 10; Gaps 1;
Qy 3 APSCAFPVQFRPQSVS-----GLSQITKSLYISNGVAAN--NKLMLSSNQITWVNV 52
Db 145 APAQALPPAGAENSNDPRVPIYDQGGPVEILPYLYLGSCNHSDDLQGLQACGITAIVLN 204
Qy 53 SVEVVNTLYEDIQVMQPVADSPNSRLCDFDPIADHIHSVEMKQGRTHLHCAAGVRS 112
Db 205 SASCPNHFEGFLHYKSIPIVEDNQMVETSAWFQEAISFTDSVKNSGGRVLVHCOAGISRSA 264
Qy 113 ALCLAYLMKYHAMSLDLAHTWTKSCRPIIRPNSGFWQLIHYEFQL 158
Db 265 TICLAYLIQSHRVLDEAFDFVKORRGVISPNSFGMLQLQLETOV 310

RESULT 15

US-09-164-193-22
; Sequence 22, Application US/09164193C
; Patent No. 6258582

GENERAL INFORMATION:

; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: NOVEL CSATPP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended)
; FILE REFERENCE: MNI-051
; CURRENT APPLICATION NUMBER: US/09/164,193C

; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-164-193-22

Query Match 21.2%; Score 208; DB 4; Length 314;
Best Local Similarity 33.1%; Pred. No. 3e-17;
Matches 55; Conservative 26; Mismatches 75; Indels 10; Gaps 1;
Qy 3 APSCAFPVQFRPQSVS-----GLSQITKSLYISNGVAAN--NKLMLSSNQITWVNV 52
Db 146 APAQALPPAGAENSNDPRVPIYDQGGPVEILPYLYLGSCNHSDDLQGLQACGITAIVLN 205
Qy 53 SVEVVNTLYEDIQVMQPVADSPNSRLCDFDPIADHIHSVEMKQGRTHLHCAAGVRS 112
Db 206 SASCPNHFEGFLHYKSIPIVEDNQMVETSAWFQEAISFTDSVKNSGGRVLVHCOAGISRSA 265
Qy 113 ALCLAYLMKYHAMSLDLAHTWTKSCRPIIRPNSGFWQLIHYEFQL 158
Db 266 TICLAYLIQSHRVLDEAFDFVKORRGVISPNSFGMLQLQLETOV 311

Search completed: April 21, 2003, 18:35:37
Job time : 23 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	818	83.4	188	11	Q8VE01	mus musculus
2	780	79.5	151	4	Q8BTG8	homo sapien
3	733.5	74.8	189	11	Q9D9D8	mus musculus
4	700	71.4	190	4	Q8H596	homo sapien
5	661	67.4	190	4	Q8F346	homo sapien
6	338	34.5	272	5	Q9XVE7	caenorhabd1
7	304	31.0	212	5	Q9VAB0	drosofophila
8	275	28.0	226	5	Q83592	caenorhabd1
9	253	25.8	353	13	Q42253	gallus gall
10	252	25.7	367	11	Q64193	rattus sp.
11	252	25.7	367	11	Q63683	rattus norv
12	252	25.7	369	13	Q91790	xenopus lae
13	249	25.4	369	13	Q90W58	xenopus lae
14	245	25.0	476	5	Q46122	drosofophila
15	240	24.5	411	4	Q13649	homo sapien
16	240	24.5	476	5	Q9VHV8	drosofophila

```

Qy 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIYK 180
Db 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIYK 180
Qy 181 EVRLMPL 188
Db 181 ETRLMPL 188

RESULT 2
Q9BTG8 ID Q9BTG8 PRELIMINARY; PRT; 151 AA.
AC Q9BTG8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 17.1 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RL Strausberg R.;
DR EMBL; BC004110; AA04110.1; -.
DR HSP; P51452; 1VHR.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 151 AA; 17111 MW; 43B60648E7023FDF CRC64;

Query Match 79.58; Score 780; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-74;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 LSSNOITWVINSVEVNTLYEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGR 100
Db 4 LSSNOITWVINSVEVNTLYEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGR 63
Qy 101 LLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFG 160
Db 64 LLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFG 123
Qy 161 KNTVHMVSSPVGMPIDPIYKEVRLMPL 188
Db 124 KNTVHMVSSPVGMPIDPIYKEVRLMPL 151

RESULT 3
Q9D9D8 ID Q9D9D8 PRELIMINARY; PRT; 189 AA.
AC Q9D9D8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1700094E07Rik protein.
GN 1700094E07RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;

Qy 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIYK 180
Db 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIYK 180
Qy 181 EVRLMPL 188
Db 181 ETRLMPL 188

RESULT 4
Q9H596 ID Q9H596 PRELIMINARY; PRT; 190 AA.
AC Q9H596
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BA386N14.1 (novel protein similar to a dual specificity phosphatase).
GN BA386N14.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Howden P.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

```

```

RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007061; BAB24847.1; -.
DR HSP; P51452; 1VHR.
DR MGD; MGI:1920797; 1700094E07Rik.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 189 AA; 21508 MW; 8EE2F7151BB5CBFA CRC64;

Query Match 74.8%; Score 733.5; DB 11; Length 189;
Best Local Similarity 73.7%; Pred. No. 1.4e-69;
Matches 137; Conservative 21; Mismatches 27; Indels 1; Gaps 1;

Qy 1 MTAPSCAFFVQ-FRQPSVGLSQTSLYISNGVAANNKMLSSNOITWVINSVEVNT 59
Db 1 MTAPSCAFFVQ-FRQPSVGLSQTSLYISNGVAANNKMLSSNOITWVINSVEVNT 60
Qy 60 LYEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGRLLHCAAGVSRSAALCLAYL 119
Db 61 FFEDIQYVQVPVADSPNSLYDFDPIADHIHSEVMKQGRLLHCAAGVSRSAALCLAYL 120
Qy 120 MKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIY 179
Db 121 MKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHVEFQFGKNTVHMVSSPVGMPIDPIY 180
Qy 180 KEVRLM 185
Db 181 KEAYLM 186

RESULT 4
Q9H596 ID Q9H596 PRELIMINARY; PRT; 190 AA.
AC Q9H596
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BA386N14.1 (novel protein similar to a dual specificity phosphatase).
GN BA386N14.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Howden P.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

```

```

DR ENBL; AL133545; CAC10195.1; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00363; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 190 AA; 21529 MW; 3E52BA31A494EE3 CRC64;

Query Match          71.4%; Score 700; DB 4; Length 190;
Best Local Similarity 68.9%; Pred. No. 5.1e-66;
Matches 131; Conservative 33; Mismatches 25; Indels 2; Gaps 1;

QY 1 MTAPSCAFFVQ--FRQPSVSGLSQITKSLYSINGVAANNKLMKSSNQITWYNVSVEVN 58
   ||| : | : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 1 MTASASSFSSSGQQVPQSIIYSFQSITRSFLSLNGVAANDKLLSNRTAIVNASVEVN 60

QY 59 TLYEDIQMVPVADSPNSRLCDFDPTADIHIHSEVKMGQRTLHCAGVSRSAALCLAY 118
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 61 VFFEIGIQIKVPVTADARDSLDFDPTADIHLIHTIDMRQGTTLHLCMAGVSRSAALCLAY 120

QY 119 LMKYHMSLLDAHWTKSCRPIIRPNSGFWQLIHFEOLFQKNTVMHVMSPVGMIPDIY 178
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 121 LMKYHMSLLDAHWTKSRPILRPNNGFWEQLINIEFKLFNNTVRMINSPVGNIPIDY 180

QY 179 EKEVRLMIPL 188
   ||| : ||| :
Db 181 EKDLRMISM 190

RESULT 5
QY9Y346 PRELIMINARY; PRT; 190 AA.
AC Q9Y346;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-JUN-2001 (TEMBLrel. 12, Last sequence update)
DE DE Hypothetical 21.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RS SEQUENCE FROM N.A.
RA Persson A.E., Sivertzen M., Lundberg J., Uhlen M.;
RT "EU-IMAGE: Full-insert length sequencing of human cDNA clones.";
RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR ENBL; AF143321; AAD3910.1; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00363; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
SQ SEQUENCE 190 AA; 21402 MW; 509A9BB85060F712 CRC64;

Query Match          67.4%; Score 661; DB 4; Length 190;
Best Local Similarity 65.8%; Pred. No. 6.8e-62;
Matches 125; Conservative 33; Mismatches 31; Indels 2; Gaps 1;

QY 1 MTAPSCAFFVQ--FRQPSVSGLSQITKSLYSINGVAANNKLMKSSNQITWYNVSVEVN 58
   ||| : | : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 1 MTASASSFSSSGQQVPQSIIYSFQSITRSFLSLNGVAANDKLLSNRTAIVNASVGSQG 60

QY 59 TLYEDIQMVPVADSPNSRLCDFDPTADIHIHSEVKMGQRTLHCAGVSRSAALCLAY 118
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 61 RILGLQTIKVPVTADARDSLDFDPTADIHLIHTIDMRQGTTLHLCMAGVSRSAALCLAY 120

QY 119 LMKYHMSLLDAHWTKSCRPIIRPNSGFWQLIHFEOLFQKNTVMHVMSPVGMIPDIY 178

```



```

OX NCBI_TaxID=10118;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=96163253; PubMed=8587253;
RX Feng L., Xia Y., Seiffert D., Wilson C.B.;
RT "Oxidative stress-inducible protein tyrosine phosphatase in
  glomerulonephritis.";
RL Kidney Int. 48:1920-1928(1995).
DR EMBL: S81478; AAB36123.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase.
SQ
SEQUENCE 367 AA; 39580 MW; EEE34CB4E04995A0 CRC64;

Query Match 25.7%; Score 252; DB 11; Length 367;
Best Local Similarity 34.3%; Pred. No. 3.1e-18;
Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps

QY 5 SCAPPVQFRPSVSGLSQITKSLYISNGVAANKMLSSNQITWLVNSVEVNTLYEDI 64
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 163 SCSTPY-----DGGPPEVILSFYLGSAVHASRKDMLDALGITALLINVSANCFHFEHY 124
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 65 QYMQVPVADSPNSRLCDFDPIADHSHVEMKGRITLLHCAAGVSRAALCLAYLMKYHA 124
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 219 QYKSIPVEDNHKADISSWFNEAIDFDSIKDAGRFEVHCQAGISRSATICLAYLMRTNR 278
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 125 MSLLDADWTWTKRCRPIRNPNSGFWEOQLIHVEFOLFGKNTVHMVSSPVGMIPD 176
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 279 VKLDEAFEFYKQRRIISPNFSFMGQLLOPESQVLAPHCSAEAGSPAMAVILD 330
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 11
Q63683
ID ID Q63683 PRELIMINARY; PRT; 367 AA.
AC AC Q63683;
DT DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Protein tyrosine phosphatase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=99396926; PubMed=10467595;
RA Qian Z., Gilbert M., Kandel E.R.;
RT "Temporal and spatial regulation of the expression of BAD2, a MAP
  kinase phosphatase, during seizure, kindling, and long-term
  potentiation.";
RL Learn. Mem. 1:180-188(1994).
DR EMBL: U02553; AAA03432.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 17:54:24 ; Search time 43 Seconds
(without alignments)
582.584 Million cell updates/sec

Title: US-09-527-376-2

Perfect score: 981

Sequence: 1 MTPASCAFFPQFRQPSVSGL.....SPVGMIPDIYEKEVRLMIPL 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	981	100.0	188	21 AAB19008	A human dual-speci
2	981	100.0	188	22 AAM39323	Human polypeptide
3	981	100.0	188	22 AAB85360	Human phosphatase
4	981	100.0	188	22 AAB73221	Human phosphatase
5	981	100.0	192	22 AAU22098	Human cardiovascular
6	981	100.0	192	22 AAM41109	Human polypeptide
7	981	100.0	192	22 AAU21820	Novel human neopla
8	921	93.9	177	22 ABB29053	Peptide #1704 enco
9	921	93.9	177	22 ABB34215	Peptide #1721 enco
10	921	93.9	177	22 AAM55008	Human brain expres

11	921	93.9	177	22 AAM67394	Human bone marrow
12	921	93.9	177	22 AAM15224	Peptide #1658 enco
13	921	93.9	177	22 AAM27685	Peptide #1722 enco
14	921	93.9	177	22 AAM02967	Human peptide enco
15	921	93.9	177	23 AAB37020	Human peptide enco
16	700	71.4	190	21 AAB36085	Human Dsp-8. Homo
17	700	71.4	190	22 AAB73220	Human phosphatase
18	697	71.0	190	22 AAM78868	Human protein SEQ
19	642.5	65.5	199	22 AAM79852	Human protein SEQ
20	511	52.1	198	20 AAY39371	Protein phosphatas
21	511	52.1	198	20 AAY28623	Human secreted pro
22	511	52.1	198	21 AAB19602	Human dual-specifi
23	511	52.1	198	21 AAY73394	Human secreted pro
24	504	51.4	198	21 AAB19620	Dual-specificity p
25	500	51.0	198	21 AAB19621	Dual-specificity p
26	431	43.9	151	22 AAM70913	Human bone marrow
27	431	43.9	151	23 ABB40709	Human peptide enco
28	304	31.0	212	22 ABB70733	Drosophila melanog
29	252	25.7	367	23 AAU84270	Human endometrial
30	252	25.7	367	23 ABB57320	Mouse ischaemic co
31	246	25.1	169	22 AAE06782	Human dual-specifi
32	246	25.1	169	22 AAB66438	Human MAP-kinase p
33	242	24.7	367	15 AAB63602	MAP-kinase-phospha
34	240	24.5	394	15 AAB63601	MAP-kinase-phospha
35	240	24.5	394	16 AAR78635	Partial MAP kinase
36	240	24.5	394	22 AAB76875	Human lung tumour
37	240	24.5	394	23 AAU85530	STY8 lung tumour p
38	240	24.5	476	22 ABB63527	Drosophila melanog
39	238	24.3	302	22 AAB76876	Human lung tumour
40	238	24.3	302	23 AAU85531	L551S lung tumour
41	238	24.3	394	22 AAB76878	Human lung tumour
42	238	24.3	394	22 AAG67448	Amino acid sequenc
43	238	24.3	394	22 AAG67627	Amino acid sequenc
44	238	24.3	394	23 AAU85533	L551S lung tumour
45	238	24.3	401	23 AAU85617	Recombinant protei

ALIGNMENTS

RESULT 1

AAB19008

ID AAB19008 standard; Protein; 188 AA.

XX AAB19008;

XX AC AAB19008;

XX DT 08-FEB-2001 (first entry)

XX DE A human dual-specificity phosphatase 2 (DSP-2).

XX KW Human; mitogen-activated protein kinase; MAP kinase; cell survival;
KW dual-specificity phosphatase; DSP-2; cell proliferation; DSP-2;
KW cell differentiation; cancer; graft-versus-host disease; allergy;
KW autoimmune disease; metabolic disease; cell cycle abnormality.

XX OS Homo sapiens.

XX PN WO200056899-A1.

XX XX 28-SEP-2000.

XX XX 22-MAR-2000; 2000WO-US07589.

XX PR 24-MAR-1999; 99US-0125957.

XX PR 16-MAR-2000; 2000US-0527376.

XX PA (CEPT-) CEPTYR INC.

XX PI Luche RM, Wei B;

XX XX WPI; 2000-638264/61.

XX DR N-PSDB; AAA96619.

XX XX

PT Novel dual-specificity mitogen activated protein (MAP) kinase
 PT phosphatase polypeptide useful in assays to identify agents that
 PT modulate the protein activity, which are used to treat cancer and
 PT autoimmune diseases -

XX Claim 1; Fig 2; 51pp; English.

XX The present sequence represents a human dual-specificity mitogen-
 CC activated protein (MAP) kinase phosphatase (DSP-2). The protein has
 CC the ability to dephosphorylate an activated MAP-kinase. The DSP-2
 CC protein is used for treatment of conditions associated with cell
 CC proliferation, cell differentiation, and cell survival. The DSP-2
 CC protein is useful for screening for modulators. Agents that modulate
 CC DSP-2 activity are useful for treating cancer, graft-versus-host
 CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
 CC cell growth, abnormal cell proliferation and cell cycle abnormalities.

XX Sequence 188 AA;

Query Match 100.0%; Score 981; DB 21; Length 188;
 Best Local Similarity 100.0%; Pred. No. 5e-110;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60
 DB 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60
 QY 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120
 DB 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120
 QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQQLIHFEFQFGKNTVHMVSSPVGMIPDIYK 180
 DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQQLIHFEFQFGKNTVHMVSSPVGMIPDIYK 180
 QY 181 EVRLMIPL 188
 DB 181 EVRLMIPL 188

RESULT 2

AAM39323
 ID AAM39323 standard; Protein; 188 AA.

XX AC AAM39323;

XX 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2469.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 23-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR N-PSDB; AAI58479.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 4; SEQ ID NO 2468; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 188 AA;

Query Match 100.0%; Score 981; DB 22; Length 188;

Best Local Similarity 100.0%; Pred. No. 5e-110;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60

DB 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60

QY 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120

DB 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120

QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQQLIHFEFQFGKNTVHMVSSPVGMIPDIYK 180

DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQQLIHFEFQFGKNTVHMVSSPVGMIPDIYK 180

QY 181 EVRLMIPL 188

DB 181 EVRLMIPL 188

RESULT 3

AAB85360

ID AAB85360 standard; Protein; 188 AA.

XX AC AAB85360;

XX 17-SEP-2001 (first entry)

XX Human phosphatase (PP) (clone ID 6205333CD1).

XX Phosphatase; PP; Alzheimer's disease; Huntington's disease; human;
 KW immune disorder; cancer; nootropic; immunomodulatory; cytostatic;
 KW neuroprotective; anticonvulsant.

XX Homo sapiens.

XX WO200153469-A2.

XX

```
PD 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US02088.
XX
XX 21-JAN-2000; 2000US-0177719.
XX
XX 28-JAN-2000; 2000US-0178988.
XX
XX 25-FEB-2000; 2000US-0184959.
XX
XX 17-MAR-2000; 2000US-0190142.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Bandman O, Tang YT, Azimzai Y, Yue H, Baughn MR, Hillman JL;
XX Lal P, Wang E, Gandhi AR, Policky JL, Mathur P;
XX
XX WPI: 2001-451905/48.
XX N-PSDB; AAH22963.
XX
XX Novel phosphatases useful for treating diseases associated with
XX decreased expression of functional phosphatases, e.g., Alzheimer's
XX disease, Huntington's disease, immune disorders, and cancers
XX
XX Claim 1; Page 95-96; 103pp; English.
XX
XX The invention provides human phosphatases (pp) and polynucleotides
XX encoding the phosphatases. The polypeptides can be expressed by standard
XX recombinant methodology. The pp are useful for treating a disease or
XX condition associated with decreased expression of functional
XX phosphatases. Compositions containing agonists or antagonists of pp may
XX be used to treat a disease associated with decreased expression or
XX overexpression of pp, respectively. Such diseases may include
XX Alzheimer's disease, Huntington's disease, immune disorders, and
XX cancers. The present sequence represents a human phosphatase.
XX
XX Sequence 188 AA;
XX
XX Query Match 100.0%; Score 981; DB 22; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 5e-110;
XX Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WTAPSCAPVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 1 WTAPSCAPVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 61 YEDIQMVQPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYLM 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 61 YEDIQMVQPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYLM 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 181 EVRLMIPL 188
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 181 EVRLMIPL 188
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 4
XX AAB73221
XX ID AAB73221 standard; Protein; 188 AA.
XX
XX AC AAB73221;
XX
XX XX 11-MAY-2001 (first entry)
XX
XX DE Human phosphatase AA915932_h.
XX
XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
XX congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
XX schizophrenia; hamartoma.
XX
XX OS Homo sapiens.
```

```
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX
XX WPI: 2001-211226/21.
XX N-PSDB; AAF63573.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
XX phosphatase related disorders such as cancer, schizophrenia, cardiac
XX dysfunction and/or vascular disorders
XX
XX Claim 6; Fig 5; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX
XX Sequence 188 AA;
XX
XX Query Match 100.0%; Score 981; DB 22; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 5e-110;
XX Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WTAPSCAPVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 1 WTAPSCAPVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 61 YEDIQMVQPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYLM 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 61 YEDIQMVQPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYLM 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 181 EVRLMIPL 188
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 181 EVRLMIPL 188
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 5
XX AAU22098
XX ID AAU22098 standard; Protein; 192 AA.
XX
XX AC AAU22098;
XX
XX XX 17-DEC-2001 (first entry)
XX
XX DE Human cardiovascular system antigen polypeptide SEQ ID No 872.
XX
XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
XX chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
XX antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
XX cerebroprotective; nootropic; antitubacterial; virucide; fungicide; cancer;
```

KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
XX Homo sapiens.
XX
XX WO200155321-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01340.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451930/48.
DR N-PSDB; AAS35372.
XX
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system -
XX
PS Claim 11; SEQ ID NO 872; 674pp; English.
XX
XX Sequences AAU21852-AAU22466 represent the cardiovascular system antigen
CC polypeptides of the invention. Cardiovascular system antigens and their
CC associated polynucleotides are useful in the diagnosis, treatment and
CC prevention of various types of disorders in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
CC can be determined by detecting the presence or absence of a mutation in a
CC cardiovascular system antigen polynucleotide. The treatable disorders
CC include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders and fungi.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi.
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
Query Match 100.0%; Score 981; DB 22; Length 192;
Best Local Similarity 100.0%; Pred. No. 5.2e-110;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTAPSCAPFVQFRQPSVSGLSQITKSLYSINGVAANNKMLSSNQITWVINVSEVWNTL 60
Db 5 MTAPSCAPFVQFRQPSVSGLSQITKSLYSINGVAANNKMLSSNQITWVINVSEVWNTL 64
Qy 61 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVENKQGRTHLHCAAGYSRAALCLAYLM 120
Db 65 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVENKQGRTHLHCAAGYSRAALCLAYLM 124
Qy 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQLHYEFQFGKNTVHMVSPVGMIPDIYEK 180
Db 125 KYHMSLLDAHTWTKSCRPIIRPNSGFWQLHYEFQFGKNTVHMVSPVGMIPDIYEK 184
RESULT 6
AA41109
ID AA41109 standard; Protein; 192 AA.
XX
AC AA41109;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6040.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI60265.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 6040; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 192 AA;
Query Match 100.0%; Score 981; DB 22; Length 192;
Best Local Similarity 100.0%; Pred. No. 5.2e-110;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTAPSCAPFVQFRQPSVSGLSQITKSLYSINGVAANNKMLSSNQITWVINVSEVWNTL 60
Db 5 MTAPSCAPFVQFRQPSVSGLSQITKSLYSINGVAANNKMLSSNQITWVINVSEVWNTL 64
Qy 61 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVENKQGRTHLHCAAGYSRAALCLAYLM 120
Db 65 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVENKQGRTHLHCAAGYSRAALCLAYLM 124
Qy 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQLHYEFQFGKNTVHMVSPVGMIPDIYEK 180
Db 125 KYHMSLLDAHTWTKSCRPIIRPNSGFWQLHYEFQFGKNTVHMVSPVGMIPDIYEK 184

QY 181 EVRLMPL 188
 Db 185 EVRLMPL 192
 RESULT 7
 AAU21820
 ID AAU21820 standard; Protein; 192 AA.
 AC AAU21820;
 XX
 XX
 DT 04-DEC-2001 (first entry)
 DE
 DE Novel human neoplastic disease associated polypeptide #253.
 XX
 KW Human; neoplastic disease associated polypeptide; cancer;
 KW hyperproliferative disorder; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.
 XX
 OS Homo sapiens.
 XX
 PN WO20015163-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01358.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465558/50.
DR N-PSDB; AAS35019.
DR
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid
PT arthritis
XX
XX
XX Claim 11; SEQ ID No 547; 687pp; English.
PS
CC The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AA021568-AA021851 represent the novel human
CC neoplastic disease associated polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
XX Sequence 192 AA;
SQ
Query Match 100.0%; Score 981; DB 22; Length 192;
Best Local Similarity 100.0%; Pred. No. 5.2e-110;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTAPSCAFPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITMWINVSEVVNTL 60
Db 5 MTAPSCAFPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITMWINVSEVVNTL 64
QY 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLM 120
Db 65 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLM 124
QY 121 KYHAMSLDADHTWTKSCRPIIRPNSGFWEQLIHYEFOLFGKNTVMVSSPVGMLPDIYEK 180
Db 125 KYHAMSLDADHTWTKSCRPIIRPNSGFWEQLIHYEFOLFGKNTVMVSSPVGMLPDIYEK 184
QY 181 EVRLMIPL 188
Db 1 EVRLMIPL 188

Db 185 EVRLMIPL 192
RESULT 8
ID ABB29053 standard; Peptide; 177 AA.
XX
XX AC ABB29053;
XX
XX DT 01-FEB-2002 (first entry)
XX
XX Peptide #1704 encoded by breast cell single exon nucleic acid probe.
DE Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00662.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 27; SEQ ID NO 12021; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
XX Sequence 177 AA;
SQ
Query Match 93.9%; Score 921; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITMWINVSEVVNTLIEDIQYMQVPV 71
Db 1 FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITMWINVSEVVNTLIEDIQYMQVPV 60

QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131
 Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120
 QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188
 Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

RESULT 9
 ABB34215
 ID ABB34215 standard; Peptide; 177 AA.
 AC ABB34215;
 XX
 DT 04-FEB-2002 (first entry)
 DE Peptide #1721 encoded by human foetal liver single exon probe.
 DE Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 OS
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver.
 XX
 PS Claim 27; SEQ ID NO 26850; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 177 AA;

Query Match 93.9%; Score 921; DB 22; Length 177;
 Best Local Similarity 100.0%; Pred. No. 8.2e-103;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVNTLYEDIQYMQVPV 71
 Db 1 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVNTLYEDIQYMQVPV 60
 QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131
 Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120

QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188
 Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

RESULT 10
 AAM55008
 ID AAM55008 standard; Protein; 177 AA.
 XX AAM55008;
 AC
 XX 05-NOV-2001 (first entry)
 DT Human brain expressed single exon probe encoded protein SEQ ID NO: 27113.
 DE Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00567.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 XX
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO: 27113; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

SQ Sequence 177 AA;

Query Match 93.9%; Score 921; DB 22; Length 177;
 Best Local Similarity 100.0%; Pred. No. 8.2e-103;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVNTLYEDIQYMQVPV 71
 Db 1 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVNTLYEDIQYMQVPV 60
 QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131
 Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120
 QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188
 Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177


```
RESULT 11
AAM67394
ID AAM67394 standard; Protein; 177 AA.
XX
AC AAM67394;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27700.
DE
DE Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
KW
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 27700; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX SQ Sequence 177 AA;
XX
Query Match 93.9%; Score 921; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVSEVNTLYEDIQYMQVPV 71
DB 1 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVSEVNTLYEDIQYMQVPV 60
QY 72 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131
DB 61 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120
QY 132 TWTKSCRPIRPNISGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 188
DB 121 TWTKSCRPIRPNISGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 177
RESULT 12
AAM15224
ID AAM15224 standard; Protein; 177 AA.
XX
AC AAM15224;
XX
DT 12-OCT-2001 (first entry)
XX
Peptide #1722 encoded by probe for measuring placental gene expression.
```

```
XX
DE Peptide #1658 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID No 20050; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI0068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 177 AA;
XX
Query Match 93.9%; Score 921; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVSEVNTLYEDIQYMQVPV 71
DB 1 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVSEVNTLYEDIQYMQVPV 60
QY 72 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131
DB 61 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120
QY 132 TWTKSCRPIRPNISGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 188
DB 121 TWTKSCRPIRPNISGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 177
RESULT 13
AAM27685
ID AAM27685 standard; Protein; 177 AA.
XX
AC AAM27685;
XX
DT 17-OCT-2001 (first entry)
XX
XX Peptide #1722 encoded by probe for measuring placental gene expression.
```

XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.

OS Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 27954; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;

CC see AAI31315-AA157546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for

CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.

XX Sequence 177 AA;

XX Query Match 93.9%; Score 921; DB 22; Length 177;

XX Best Local Similarity 100.0%; Pred. No. 8.2e-103;

XX Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FRQPSVGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNTLYEDIQVMQPV 71

DB 1 FRQPSVGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNTLYEDIQVMQPV 60

QY 72 ADSPNSRLCDFDPTADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 131

DB 61 ADSPNSRLCDFDPTADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 120

QY 132 TWTKSCRPIIRPNISGFWQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 188

DB 121 TWTKSCRPIIRPNISGFWQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 177

RESULT 14

AA02967

ID AA02967 standard; Protein; 177 AA.

XX AA02967;

XX 09-OCT-2001 (first entry)

XX Peptide #1649 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

PN

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US00661.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression

PT in a human breast -

XX Claim 27; SEQ ID No 11707; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes

CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for measuring human gene expression in

CC a human breast sample, where the probe hybridizes at high stringency to a

CC nucleic acid expressed in the human breast. The probes are useful for

CC predicting, diagnosing, grading, staging, monitoring and prognosing

CC diseases of the human breast, particularly those diseases with polygenic

CC aetiology. The diseases include: breast cancer, disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 177 AA;

XX Query Match 93.9%; Score 921; DB 22; Length 177;

XX Best Local Similarity 100.0%; Pred. No. 8.2e-103;

XX Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FRQPSVGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNTLYEDIQVMQPV 71

DB 1 FRQPSVGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNTLYEDIQVMQPV 60

QY 72 ADSPNSRLCDFDPTADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 131

DB 61 ADSPNSRLCDFDPTADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 120

QY 132 TWTKSCRPIIRPNISGFWQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 188

DB 121 TWTKSCRPIIRPNISGFWQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 177

RESULT 15

ABG37020

ID ABG37020 standard; Peptide; 177 AA.

XX ABG37020;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 26685.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
primary ciliary dyskinesia; pulmonary hypertension;
hyaline membrane disease.

Homo sapiens.

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-

20 MAY 2000; 2000US-207430E;
30-JUN-2000; 2000US-0608408.

03 AUG 2000; 2000US-0032300;
21-SEP-2000; 2000US-234687P;

27 SEP 2000; 2000GB-0024263;
04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS I

Penn SG, Hanzel DK, Chen W

WPI; 2002-114183/15.

Spatially-addressable set of si

SECRET

ON 12 DEC 1977

The invention relates to a spatially addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of

nucleic acid expressed in the human lung;

a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 177 AA;

	Query Match	93.9%;	Score 921;	DB 23;	Length 177;
	Best Local Similarity	100.0%;	Prod. No. 8.2e-103;		
	Matches 177;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	12	FRQPSVSGLSQITKSLYISNGVAANNKMLMLSSNQITMVINVSVEVNTLYEDIQYMQVPV	71		
Db					
Db	1	FRQPSVSGLSQITKSLYISNGVAANNKMLMLSSNQITMVINVSVEVNTLYEDIQYMQVPV	60		
Qy	72	ADSPNSRLCDDFFDIADHIIHSVENMKQGRTHLLHCAAGYSRSGAALCLAYLMKYHAMSLLDAH	131		
Db	61	ADSPNSRLCDDFFDIADHIIHSVENMKQGRTHLLHCAAGYSRSGAALCLAYLMKYHAMSLLDAH	120		
Qy	132	TWTKSCRPRIIRPNSGFWEOLIHYEFQFGKNTVHMVSSPGMIPDIIYEKEVRLMIPL	188		
Db	121	TWTKSCRPRIIRPNSGFWEOLIHYEFQFGKNTVHMVSSPGMIPDIIYEKEVRLMIPL	177		

Search completed: April 21, 2003, 18:32:34
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 18:35:04 ; Search time 114 seconds
(without alignments)
124.637 Million cell updates/sec

Title: US-09-527-376-2
Perfect score: 981
Sequence: 1 MTAPSCAFFVQFRQPSVGL.....SPVGMIPDIYEKVRMLMPL 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	981	100.0	192	9	US-10-091-504-872
2	981	100.0	192	10	US-09-764-869-872
3	921	93.9	177	10	US-09-864-761-42750
4	700	71.4	190	12	US-10-044-205A-44
5	697	71.0	190	12	US-10-044-205A-42
6	431	43.9	151	10	US-09-864-761-46915
7	252	25.7	367	10	US-09-919-497-60
8	246	25.1	154	9	US-09-955-732-8
9	246	25.1	154	10	US-09-964-277-9
10	246	25.1	169	10	US-09-775-925-28
11	240	24.5	169	10	US-09-847-519A-13
12	240	24.5	394	9	US-09-736-457-805
13	240	24.5	394	9	US-09-902-941-805
14	240	24.5	394	9	US-09-849-626-805
15	240	24.5	394	9	US-10-017-754-805
16	238	24.3	302	9	US-09-736-457-806
17	238	24.3	302	9	US-09-902-941-806
18	238	24.3	302	9	US-09-849-626-806
19	238	24.3	302	9	US-10-017-754-806

20	238	24.3	394	9	US-09-736-457-827	Sequence 827, App
21	238	24.3	394	9	US-09-902-941-827	Sequence 827, App
22	238	24.3	394	9	US-09-849-626-827	Sequence 827, App
23	238	24.3	394	9	US-10-017-754-827	Sequence 827, App
24	238	24.3	401	9	US-09-902-941-1917	Sequence 1917, App
25	238	24.3	401	9	US-09-849-626-1917	Sequence 1917, App
26	238	24.3	401	9	US-10-017-754-1917	Sequence 1917, App
27	230	23.4	235	10	US-09-963-204-2	Sequence 2, Appli
28	230	23.4	289	10	US-09-963-204-12	Sequence 12, Appl
29	230	23.4	329	10	US-09-963-204-22	Sequence 22, Appl
30	227	23.1	154	9	US-09-955-732-9	Sequence 9, Appli
31	227	23.1	154	10	US-09-964-277-10	Sequence 10, Appl
32	227	23.1	169	10	US-09-775-925-29	Sequence 29, Appl
33	227	23.1	169	10	US-09-847-519A-14	Sequence 14, Appl
34	221	22.5	154	9	US-09-955-732-10	Sequence 10, Appl
35	221	22.5	154	10	US-09-964-277-11	Sequence 11, Appl
36	221	22.5	171	10	US-09-775-925-30	Sequence 30, Appl
37	221	22.5	171	10	US-09-847-519A-15	Sequence 15, Appl
38	220	22.4	173	10	US-09-816-494-7	Sequence 7, Appli
39	220	22.4	173	10	US-09-815-419-4	Sequence 4, Appli
40	219.5	22.4	170	10	US-09-775-925-26	Sequence 26, Appl
41	219.5	22.4	170	10	US-09-847-519A-11	Sequence 11, Appl
42	217	22.1	217	10	US-09-815-419-2	Sequence 2, Appli
43	213.5	21.8	172	10	US-09-816-494-8	Sequence 8, Appli
44	213.5	21.8	172	10	US-09-815-419-5	Sequence 5, Appli
45	211	21.5	155	9	US-09-955-732-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-10-091-504-872
; Sequence 872, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 872
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-504-872

Query Match 100.0%; Score 981; DB 9; Length 192;
Best Local Similarity 100.0%; Pred. No. 9.3e-101;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTAPSCAFFVQFRQPSVGLSQITKSLYTSNGVAANKMLSSNQITWINSVEVVNTL	60
DB	5	MTAPSCAFFVQFRQPSVGLSQITKSLYTSNGVAANKMLSSNQITWINSVEVVNTL	64
QY	61	YEDIQMVQPVADSPNSRLCDFDPDIADHHSVEMKQGRLLHCAAGVSRAALCLAYLM	120
DB	65	YEDIQMVQPVADSPNSRLCDFDPDIADHHSVEMKQGRLLHCAAGVSRAALCLAYLM	124
QY	121	KYHMSLLDAHTWTKSCRIIPRNSGFWQLIHYEQLFQKNTVHMVSSPVGMPIDYK	180
DB	125	KYHMSLLDAHTWTKSCRIIPRNSGFWQLIHYEQLFQKNTVHMVSSPVGMPIDYK	184
QY	181	EVRLMIPL 188	
DB	185	EVRLMIPL 192	
RESULT 2			
US-09-764-869-872			

[illegible]

QY		65	OYMQVPVADSPNRLCDFDPDIADHHTSVEMKOGRTLLHCAAGVSRSAALCLAYLWKHYA	124
	I I : I I :		I I : I I :	
Db		72	OYSIPVEDNHKADISSWFNEADFDSIKNAGRPFVHCQAGISRSATICLAYLMRTNR	131
	I I : I I :		I I : I I :	
QY		125	MSLDHTWTKSCRPIIRPNSGWEQLIHYEFOI	158
	I I : I I :		I I : I I :	
Db		132	VKLDEAFEFVKQRSSIISPNSFMGLOLFESOV	165
	I I : I I :		I I : I I :	

```

RESULT 11
US-09-847-519A-13
; Sequence 13, Application US/09847519A
; Patent No. US20020102693A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.422
; CURRENT APPLICATION NUMBER: US/09/847,519A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-519A-13

```

	Query Match	25.1%	Score 246;	DB 10;	Length 169;
	Best Local Similarity	36.4%;	Pred. No. 1.7e-19;		
	Matches	56;	Conservative 27;	Mismatches 67;	Indels 4; Gaps 1;
Qy	5	S C A F P V Q R O P S Y G L S Q I T K S L I N S G V A N N K M L S S N Q I T W V I N V S E V V A T L E D I	64		
Dd	16	S C S P L Y - - - - D O G G P V E I L F P L Y G A Y H A S R K D M L D A L G I T A L I N V S A N C P H F E G H Y	71		
Qy	65	Y Q M V O P A D S P N S R L C D F F O P I A D H I S V E M K Q C R T L L H C A A G Y S R S A A L C L A Y L M K V H A	124		
Dd	72	Y Q K S I P E D N H K A D I S W F N E A I D F I S K N A G E R V F V H C O A G I S R S A T I C L A Y L M R T N R	131		
Qy	125	M S L L D A H T W T K S C R P I I R N P S G F W E Q L I H Y F Q L	158		
Dd	132	V K L D E A F E F Y K O R S I I S P N F S F M G O L L O F E S O V	165		

```

RESULT 12
US-09-736-457-805
; Sequence 805, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-805

```

[illegible]

```

RESULT 13
US-09-902-941-805
; Sequence 805, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF LUNG CANCER
; TITLE OF INVENTION: COMPOSITIONS AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-805

```

Query Match	24.5%	Score 240;	DB 9;	Length 394;
Best Local Similarity	33.9%	Pred. No. 2.5e-18;		
Matches	57;	Conservative 32;	Mismatches 75;	Indels 4;
Gaps				

QY	5	SCAPVQFRQPSVGLSQITKSLYSINGVAANKMLSSNQITWIVNVSEVVNTLYEDI	64
		: : : : : : : : : : : : : :	
Db	185	SCGTPLHDQE----GPVEILPFLVLSGAYHAARDMLDALGITALLNVSDCPNHFEGHY	240
		: : : : : : : : : : : : : :	
QY	65	QYMOVVADSPNSRLCDDFFQPADHHSVEMQKQETLLHCAAGVSRAALCLAYLMKYHA	124
		: : : : : : : : : : : : : :	
Db	241	QYKICIPVEDNHKADJISWFMEAEIYDAVDCRGRLVHCQAGISRATCCLAYLMKKR	300
		: : : : : : : : : : : : : :	
QY	125	MSLDAHTWTKSCRPIIRPNSGFWEQLIHFEQFLFGNTVHMVSSPGV	172
		: : : : : : : : : : : : :	
Db	301	VRLEAEFEFKRRSILSPNFSFGQLIQFESQVLAATSCAAEAASPGS	348
		: : : : : : : : : : : : :	

```

RESULT 14
US-9-849-626-805
; Sequence 805, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong

```

```
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-805

Query Match      24.5%; Score 240; DB 9; Length 394;
Best Local Similarity 33.9%; Pred. No. 2.5e-18;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

Qy 5 SCAPPVQRPQSVGLSQITKSLYISNGVAANNKMLSSNOITMVINVSVVWNTLYEDI 64
   || || : : : : || : : : : || : : : : || : : : : ||
Db 185 SCGTPLDHQE---GPVEILPFLYLGSAHYAARDMLDALGITALLNVSSDCPNHFEGHY 240

Qy 65 QYMQVPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYILMKYHA 124
   || || : : : : || : : : : || : : : : || : : : : ||
Db 241 QYKICPVEDNHKADISSWFMEAEIYDAVKDCRGRVLVHCQAGISRATICLAYILMKKR 300

Qy 125 MSLDHAHTWTKSCRPIIRPNISGFWQELIHYEFQFGKNTVHMVSSPVG 172
   || || : : : : || : : : : || : : : : || : : : : ||
Db 301 VRLEAEFEFVKQRRIISPNFSFMGQLLOFESQVLATSCAAEAASPSG 348

RESULT 15
US-10-017-754-805
; Sequence 805, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-754-805

Query Match      24.5%; Score 240; DB 9; Length 394;
Best Local Similarity 33.9%; Pred. No. 2.5e-18;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

Qy 5 SCAPPVQRPQSVGLSQITKSLYISNGVAANNKMLSSNOITMVINVSVVWNTLYEDI 64
   || || : : : : || : : : : || : : : : || : : : : ||
Db 185 SCGTPLDHQE---GPVEILPFLYLGSAHYAARDMLDALGITALLNVSSDCPNHFEGHY 240

Qy 65 QYMQVPVADSPNSRLCDFDPIADHHSVENKQGRTHLHCAAGVSRSAALCLAYILMKYHA 124
   || || : : : : || : : : : || : : : : || : : : : ||
Db 241 QYKICPVEDNHKADISSWFMEAEIYDAVKDCRGRVLVHCQAGISRATICLAYILMKKR 300
```

```
Qy 125 MSLDHAHTWTKSCRPIIRPNISGFWQELIHYEFQFGKNTVHMVSSPVG 172
   || || : : : : || : : : : || : : : : || : : : : ||
Db 301 VRLEAEFEFVKQRRIISPNFSFMGQLLOFESQVLATSCAAEAASPSG 348

Search completed: April 21, 2003, 18:41:45
Job time : 115 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 18:31:54 ; Search time 27 Seconds
(without alignments)
669.380 Million cell updates/sec

Title: US-09-527-376-2
 Perfect score: 981
 Sequence: 1 MTAPSCAFPVQFRQPSVSL.....SPVGMIPDIYEKEVRLMIPL 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :      PIR_73:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	338	34.5	272	2	T18915	hypothetical prote
2	275	28.0	226	2	T21380	hypothetical prote
3	252	25.7	367	1	S29090	dual specificity p
4	252	25.7	367	1	S24411	dual specificity p
5	252	25.7	367	2	S52265	dual specificity p
6	238	24.3	303	2	T46405	hypothetical prote
7	236	24.1	393	2	A56947	dual specificity p
8	236	24.1	394	2	A56115	dual specificity p
9	223	23.8	619	2	T15969	hypothetical prote
10	231.5	23.6	186	2	T16056	hypothetical prote
11	223.5	22.8	365	2	T32494	hypothetical prote
12	221	22.5	384	1	I38890	dual specificity p
13	208	21.2	314	2	B57126	dual specificity p
14	197	20.1	314	1	A57126	dual specificity p
15	180	18.3	223	2	I49365	protein tyrosine p
16	170	17.3	142	2	T30074	dual specificity p
17	165.5	16.9	292	2	A41012	hypothetical prote
18	164.5	16.8	330	2	T39698	protein tyrosine p
19	162.5	16.6	185	1	A47196	dual specificity p
20	162	16.5	866	2	F88481	dual specificity p
21	160	16.3	209	1	S48459	probable dual spec
22	158	16.1	771	2	T47666	phosphatase-like p
23	157	16.0	205	2	I49364	protein tyrosine p
24	155	15.8	489	1	S58725	dual specificity p
25	151	15.4	171	1	B47452	dual specificity p
26	147	15.0	171	1	Q0VZH1	dual specificity p
27	147	15.0	171	1	A42514	dual specificity p
28	144	14.7	171	1	I36845	dual specificity p
29	144	14.7	171	2	T28522	probable dual spec

30	144	14.7	171	2	B72151	J1L protein - vari
31	141.5	14.4	169	2	T30684	probable dual spec
32	141.5	14.4	283	2	G3458	probable protein p
33	139.5	14.2	276	2	T48906	protein-tyrosine-p
34	137	14.0	272	2	T19418	hypothetical prote
35	137	14.0	807	1	S44538	probable protein-t
36	135	13.8	204	2	T17802	hypothetical prote
37	127	12.9	278	2	T39517	dual-specificity M
38	123.5	12.6	600	2	T18446	hypothetical prote
39	119	12.1	580	2	T18439	hypothetical prote
40	117	11.9	364	1	S31304	protein-tyrosine-p
41	109	11.1	597	1	S43743	probable dual spec
42	107	10.9	205	2	G95860	conserved hypotet
43	106	10.8	347	2	T49097	hypothetical prote
44	100.5	10.2	681	2	B88158	protein Cl7G10.4c
45	100.5	10.2	708	2	T34098	hypothetical prote

ALIGNMENTS

RESULT 1
T18915
hypothetical protein C04F12.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18915
R:Uloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19044
A:Accession: T18915
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: EMBL:Z81461; PIDN:CAB03837.1; GSPDB:GNO0019; CESP:C04F12.8
A:Experimental source: clone C04F12
C:Genetics:
A:Gene: CESP:C04F12.8
A:Map position: 1
A:Introns: 53/1; 84/3; 204/2

	Query Match	34.5%;	Score 338; DB 2; Length 272; Best Local Similarity 40.7%; Pred. No. 3.2e-25; Matches 72; Conservative 40; Mismatches 59; Indels 6; Gaps 4;
y	13 RQPSVSG-LSQITKSLYISN-GVAANNKLSSNOITWVINVSVEVVWTLYEDIQMVP 70 <div> : : :</div>		
b	3 RNVLGQGMSINDHLYLSGAGLVLPDK--IKQRINMIVNATTEEPSTYMGGVDTKIR 60 <div> : : :</div>		
y	71 VADSPNSRCDFFDPIADIHSVENMKQGTLLRLCAAGYSRSAAALCIAYLMKYHAMSLDDA 130 <div> : : :</div>		
b	61 IEDHPYARLNEHPDVADVADRKNRVKGCKTLVHCMAGYSRASLIWMYLVKHEHTLRQA 120 <div> : : :</div>		
y	131 HTWKSCRPRIIPRNSEGFWEQLIHFEOLFGRKNTVMHWSPV--GMIPDIYEKEVRML 185 <div>: : :</div>		
:	121 YHYKAARPRIIPRNSEGWOMDYEKRLRGTSASKRMVOPTCEMDPDPVIADDAIDRRM 177 <div>: : :</div>		

RESULT 2
T21380
hypothetical protein F26A3.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21380
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19415
A:Accession: T21380
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <WIL>
A:Cross-references: EMBL:Z78419; PIDN:CAB01700.1; GSPDB:GN00019; CBSP:F26A3
A:Experimental source: clone F26A3

C;Genetics:

A;Gene: CESP:F26A3.4

A;Map position: 1

A;Introns: 117/2; 150/3; 186/3

Query Match 28.0%; Score 275; DB 2; Length 226;

Best Local Similarity 36.8%; Pred. No. 3.3e-19;

Matches 68; Conservative 42; Mismatches 63; Indels 12; Gaps 7;

Qy 10 VQFR-QPSVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNN-TLYEDIQYM 67

Db 3 LSFERNPEYAMSEIVPGLFIC-GVSALSKDKMKKHITHIINATTEVPNLSRSLGDIORT 61

Qy 68 QVPVADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHMSL 127

Db 62 KLWLEDTPQTYIYPHLEQLSDIOIALIADGGKVLVHCVAGVSRASICLAFLLKYRCNL 121

Qy 128 LDAHTWTWTKSCRPIIRPNISGFWEQLIHFEQLFGKN--TVHMV---SSPVGMIPDIYEK 182

Db 122 REAYHLMKSRSMVRPNLGFWRQLIAYE-QNVKENAGSVRLVRDEAQPEQLLPDVI---L 177

Qy 183 RLMP 187

Db 178 NIAIP 182

RESULT 3

S29090

N;Alternate names: dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human

C;Species: Homo sapiens (man)

C;Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

A;Accession: S29090; A53052

R;Keyse, S.M.; Emelie, E.A.

Nature 359, 644-647, 1992

A;Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine

A;Reference number: S29090; MUID:93024952; PMID:1406996

A;Accession: S29090

A;Molecule type: mRNA

A;Residues: 1-367 <KEY>

A;Cross-references: EMBL:X68277; NID:g29980; PIDN:CAA48338.1; PID:g29981

R;Wak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.

J. Biol. Chem. 269, 3596-3604, 1994

A;Title: Isolation and characterization of a human dual specificity protein-tyrosine pho

A;Reference number: A53052; MUID:94148864; PMID:8106404

A;Accession: A53052

A;Molecule type: DNA

A;Residues: 1-367 <KWA>

A;Experimental source: leukocyte

A;Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,

C;Genetics:

A;Gene: GDB:DUSP1; PTPN10

A;Cross-references: GDB:136197; OMIM:600714

A;Map position: 5q34-5q34

C;Superfamily: dual specificity phosphoprotein phosphatase 1; vhl-type dual specificity

C;Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced

F;181-312/Domain: vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F;258/Active site: Cys (phosphocysteine intermediate) #status predicted

F;264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 25.7%; Score 252; DB 1; Length 367;

Best Local Similarity 34.3%; Pred. No. 1.1e-16;

Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

Qy 5 SCAPFVQRPQPSVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNN-TLYEDI 64

Db 163 SCSTPLY----DQGGPVEILPFLYLGAYHASRKMDLDALGITALINVSANCPNHFEGHY 218

Qy 65 QYQMPVADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHA 124

Db 219 QYKSIPIVEDNHKADISSWFNEAIDFISIKNAGGRVVFHCVQAGISRSATICLAYLMRTNR 278

Qy 125 MSLDAHTWTWTKSCRPIIRPNISGFWEQLIHFEQLFGKNTVHMVSSPVGMIPD 176

Db 279 VKLDEAFEFVKQRSSIIISPNFSGQLLQFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 5

S52265

N;Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase,

C;Species: Rattus norvegicus (Norway rat)

C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999

R;Muda, M.; Schlegel, W.; Arkinstall, S.

Submitted to the EMBL Data Library, January 1995

A;Description: Pathways regulating CL100 gene expression in pituitary cells.

A;Reference number: S52265

A;Accession: S52265

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-367 <MUD>

A;Cross-references: EMBL:X84004; NID:g642264; PIDN:CAA58828.1; PID:g642265

C;Superfamily: dual specificity phosphoprotein phosphatase 1; vhl-type dual specificity

C;Keywords: phosphoprotein; phosphoric monoester hydrolase

F;181-312/Domain: vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F;258/Active site: Cys (phosphocysteine intermediate) #status predicted

Db 279 VKLDEAFEFVKQRSSIIISPNFSGQLLQFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 4

S24411

N;Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type

C;Species: Mus musculus (house mouse)

C;Date: 19-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999

C;Accession: A54681; S24411

R;Noguchi, T.; Metz, R.; Chen, L.; Mattel, M.G.; Carrasco, D.; Bravo, R.

Mol. Cell. Biol. 13, 5195-5205, 1993

A;Title: Structure, mapping, and expression of erp, a growth factor-inducible gene en

A;Reference number: A54681; MUID:93360956; PMID:8355678

A;Accession: A54681

A;Molecule type: DNA

A;Residues: 1-367 <NOG>

A;Cross-references: GB:S64851; NID:g409976; PIDN:AAB27882.1; PID:g409977

R;Charles, C.H.; Abler, A.S.; Lau, L.F.

Oncogene 7, 187-190, 1992

A;Title: cDNA sequence of a growth factor-inducible immediate early gene and characte

A;Reference number: S24411; MUID:92158357; PMID:1741163

A;Accession: S24411

A;Molecule type: mRNA

A;Residues: 1-367 <CHA>

A;Cross-references: EMBL:X61940; NID:g49735; PIDN:CAA43944.1; PID:g49736

C;Genetics:

A;Gene: erp

A;Introns: 123/1; 172/1; 245/1

C;Superfamily: dual specificity phosphoprotein phosphatase 1; vhl-type dual specific

C;Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase

F;181-312/Domain: vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F;258/Active site: Cys (phosphocysteine intermediate) #status predicted

F;264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 25.7%; Score 252; DB 1; Length 367;

Best Local Similarity 34.3%; Pred. No. 1.1e-16;

Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

Qy 5 SCAPFVQRPQPSVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVNN-TLYEDI 64

Db 163 SCSTPLY----DQGGPVEILSFLYLGAYHASRKMDLDALGITALINVSANCPNHFEGHY 218

Qy 65 QYQMPVADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHA 124

Db 219 QYKSIPIVEDNHKADISSWFNEAIDFISIKDAGGRVVFHCVQAGISRSATICLAYLMRTNR 278

Qy 125 MSLDAHTWTWTKSCRPIIRPNISGFWEQLIHFEQLFGKNTVHMVSSPVGMIPD 176

Db 279 VKLDEAFEFVKQRSSIIISPNFSGQLLQFESQVLAPHCSAEAGSPAMAVLD 330

F:202-333/Domain: VH1-type dual specificity phosphoprotein phosphatase hom
F:279/Active site: Cys (phosphocysteine intermediate) #status predicted

A:Reference number: Z18439
A:Accession: T15969

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-619 <CH>
A:Cross-references: EMBL:U23178; NID:g726421; PID:g726422; PIDN:AAC46719.1; CESP:F08B1.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F08B1.1
A:Map position: 57/1, 99/1, 160/3, 252/2, 549/3
A:Introns: 22/2, 57/1, 99/1, 160/3, 252/2, 549/3

Query Match 23.8%; Score 233; DB 2; Length 619;
Best Local Similarity 33.3%; Pred. No. 1.6e-14;
Matches 48; Conservative 35; Mismatches 59; Indels 2; Gaps 1;
QY 14 QPSVSGLSQITKSLYSINGVAANNKMLSSNOITWVINSVEVNT--LYEDIOYMQVPV 71
DB 132 QPTGCGIILTPNIIYNGSQISDLSDTDLADISVIVNLSMFCPKSVCKEDKFMRI 191
QY 72 ADSPNSRLCDFDPIADHSHVEMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 131
DB 192 NDSYQEKLSPPPMAYEFLEKRRAGKCLHCLAGISRSPTLAISIMRYMKMGSDDAY 251
QY 132 TWTSCRPILIRPNSGFWEOLHYE 155
DB 252 RYVKRRPESISPNNFMGQLLEYE 275

RESULT 10

T16056
hypothetical protein F13D11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001
C:Accession: T16056
R:Fulton, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F13D11.
A:Reference number: 569020
A:Accession: T16056
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-186 <F>
A:Cross-references: EMBL:U40939; NID:g1072175; PID:g1072176; PIDN:AAA81700.1; CESP:F13D11.3
C:Genetics:
A:Gene: CESP:F13D11.3
A:Introns: 30/3, 57/3, 85/3, 125/3, 172/3

Query Match 23.6%; Score 231.5; DB 2; Length 186;
Best Local Similarity 31.4%; Pred. No. 4.3e-15;
Matches 55; Conservative 36; Mismatches 61; Indels 23; Gaps 4;

QY 17 VSGLSQITKSLY----ISNGVAANNKMLSSNOITWVINSVEVNTLYEDIOYMQVPVA 72
DB 20 LAGYGCITPSSLKQYNITHGVDCN--LTKKPI-----KGLDRIEVPVD 61
QY 73 DSPNSRLCDFDPIADHSHVEMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLDDAH 132
DB 62 DNTLAKITQYEPVVKYIEDAKQQGHNTVIYCAAGVSRSAATITVYLMVTENLSLEAYL 121
QY 133 WTKSCRPILIRPNSGFWEOLHYEFOLFGKNTVHMVSSPVGMIPIDIYEKVRLEMI 186
DB 122 QVNRVPIISPNIGFWRMIDFQKRNAGSVELISGRMARPVPSVILRRVTII 176

RESULT 11

T32494
hypothetical protein C05B10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32494
R:Geisel, C.; Wamsley, P.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C05B10.
A:Reference number: 221178
A:Accession: T32494

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <GT>
A:Cross-references: EMBL:AF036585; PIDN:AAB8308.1; GSPDB:GN00022; CESP:C05B10.1
A:Experimental source: strain Bristol N2; clone C05B10
C:Genetics:
A:Gene: CESP:C05B10.1
A:Map position: 4
A:Introns: 22/2, 46/2, 74/3, 163/2, 204/2, 248/1, 287/2

Query Match 22.8%; Score 223.5; DB 2; Length 365;
Best Local Similarity 33.3%; Pred. No. 6.4e-14;
Matches 56; Conservative 25; Mismatches 68; Indels 19; Gaps 4;
QY 8 FPVQFRQPSVSGLSQITKSLYSINGVAANNKMLSSNOITWVINSVEVNTLYED--IQ 65
DB 180 FPV-----KLTNFLYLGNATKRDVLYKYSISHVINVTSLNPTFEEDPNMR 228
QY 66 YMQVPVADSPNSRLCDFDPIADHSHVEMKQGRTHLHCAAGVSRSAALCLAYLMKYHAM 125
DB 229 YLRISADDNASHNLTKFFPEAISFIDDARRNDSACLHCLAGISRSVTICLAYLMKTEMC 288
QY 126 SLDDAHTWTKSCRPIIRPNSGFWEOLHYEFOLFGKNTVHMVSSPVGMI 173
DB 289 TLDSAYEWMQVRNASIAPNFMGQLTQYE-KMLGLN-----SNRQGV 330

RESULT 12

I38890
dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N:Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 11-Jun-1999
C:Accession: I38890; A55313
R:Kwak, S.P.; Dixon, J.E.
J. Biol. Chem. 270, 1156-1160, 1995
A:Title: Multiple dual specificity protein tyrosine phosphatases are expressed and re
A:Reference number: A55432; MUID:95138103; PMID:7836374
A:Accession: I38890
A:Molecule type: mRNA
A:Residues: 1-384 <RES>
A:Cross-references: EMBL:U16996; NID:g642012; PIDN:AAB06261.1; PID:g642013
R:Experimental source: placenta
R:Ishibashi, T.; Bottaro, D.P.; Michielli, P.; Kelley, C.A.; Aaronson, S.A.
J. Biol. Chem. 269, 29897-29902, 1994
A:Title: A novel dual specificity phosphatase induced by serum stimulation and heat s
A:Reference number: A55313; MUID:95050849; PMID:7961985
A:Accession: A55313
A:Molecule type: mRNA
A:Residues: 1-8,'GHV',12-70,'R',72-104,'F',107-362,'RCLPTQSSSAAELMQRPNPARTGMESAQPO
A:Cross-references: GB:U15932; NID:g606971; PIDN:AAA64693.1; PID:g606972
A:Experimental source: mammary epithelial cells
C:Genetics:
A:Gene: GDB:DUSP5
A:Cross-references: GDB:385447
A:Map position: 10q25-10q25
C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specific
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:186-317/Domain: VHI-type dual specificity phosphoprotein phosphatase homology <VHI>
F:263/Active site: Cys (phosphocysteine intermediate) #status predicted
F:269/Binding site: substrate phosphate (Arg) #status predicted

Query Match 22.5%; Score 221; DB 1; Length 384;
Best Local Similarity 34.3%; Pred. No. 1.2e-13;
Matches 48; Conservative 24; Mismatches 68; Indels 0; Gaps 0;
QY 19 GLSQITKSLYSINGVAANNKMLSSNOITWVINSVEVNTLYEDIOYMQVPVADPNRS 78
DB 178 GPVEILPPLYLGSAYHASKCEFLANLHITALLNVSRRSEACMTLHYKWIPIVEDSHAD 237
QY 79 LCDFFDPIADHSHVEMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLDDAHTWTKSCR 138
DB 238 ISSHQEAIDFIDCVREKGGKVLVHCEAGISRSPTICMAYLMKTKQFLKEAFDIKORR 297

F;257/Active site: Cys (phosphocysteine intermediate) #status predicted
F;263/Binding site: substrate phosphate (Arq) #status predicted

Search completed: April 21, 2003, 18:34:58
Job time : 29 secs

Search completed: April 21, 2003, 18:34:58
Job time : 29 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 16:59:44 ; Search time 46 seconds
(without alignments)
5560.183 Million cell updates/sec

Title: US-09-527-376-1

Perfect score: 834

Sequence: 1 ctttctctgtatttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	177.6	21.3	1729	4	US-09-045-973-6
2	87	10.4	1238	2	US-08-530-290-11
3	81.2	9.7	1987	2	US-08-990-379-1
4	81.2	9.7	1993	2	US-08-990-379-2
5	61.6	7.4	1691	3	US-09-013-881-12
6	58.8	7.1	789	4	US-09-164-193-6
7	58.8	7.1	789	4	US-09-221-448A-6
8	58.8	7.1	1016	4	US-09-164-193-4
9	58.8	7.1	1016	4	US-09-221-448A-4
10	52.2	6.3	531	4	US-09-704-139-3
11	52.2	6.3	594	4	US-09-163-833-3
12	52.2	6.3	912	4	US-09-163-833-1
13	52.2	6.3	1390	4	US-09-704-139-1
14	45	5.4	216	2	US-08-530-290-9
15	38.2	4.6	970	4	US-09-071-035-183
16	38.2	4.6	1110	4	US-09-071-035-181
17	37.4	4.5	1293	4	US-09-078-294-10
18	37	4.4	7218	1	US-08-232-463-14
19	35.8	4.3	599	4	US-09-078-294-23
20	34.4	4.1	289	4	US-09-007-005-17
21	34.4	4.1	289	4	US-09-244-796-17
22	34.4	4.1	677	2	US-08-666-405-9
23	34.4	4.1	882	2	US-08-666-405-27
24	34.4	4.1	1730	3	US-09-222-817-11
25	34.4	4.1	1730	3	US-09-222-817-13
26	34.4	4.1	1730	4	US-09-222-786-11
27	34.4	4.1	1730	4	US-09-222-786-13

Sequence 15, Appl
Sequence 37, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 134, App
Sequence 1, Appl
Sequence 2, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 611, App
Patent No. 5447867
Sequence 2, Appl
Sequence 116, App

28 34.2 4.1 17327 1 US-07-906-871-15
29 33.2 4.0 937 3 US-09-248-335-37
30 32.6 3.9 5798 2 US-08-483-101-1
31 32.4 3.9 861 1 US-07-988-273-1
32 32.4 3.9 861 5 PCT-US93-12019-1
33 32.4 3.9 1011 3 US-09-095-163-1
34 32.4 3.9 1029 4 US-09-485-648-5
35 32.4 3.9 1029 4 US-09-503-565-5
36 32.4 3.9 1029 4 US-09-485-649-5
37 32.4 3.9 1345 1 US-08-525-654A-134
38 32.4 3.9 3332 4 US-09-448-806C-1
39 32.4 3.9 4403765 4 US-09-103-840A-2
40 32 3.8 1393 1 US-08-174-467-18
41 32 3.8 1393 3 US-08-452-071-18
42 31.8 3.8 1023 4 US-09-134-001C-611
43 31.8 3.8 1636 6 5447867-2
44 31.8 3.8 2235 4 US-09-153-804-2
45 31.8 3.8 3079 4 US-09-643-597-116

ALIGNMENTS

RESULT 1
US-09-045-973-6
; Sequence 6, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,973
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0491 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1729 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT16
; CLONE: 3041794
US-09-045-973-6

[illegible]

US-09-163-833-1

; Sequence 1, Application US/09163833

; Patent No. 6268135

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: No. 6268135el Phospholipase Molecule and Uses Therefor

; FILE REFERENCE: mni-059

; CURRENT APPLICATION NUMBER: US/09/163,833

; CURRENT FILING DATE: 1998-09-30

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 912

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (52)..(645)

; US-09-163-833-1

Query Match 6.3%; Score 52.2; DB 4; Length 912;
Best Local Similarity 53.9%; Pred. No. 3.6e-06;
Matches 130; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

QY 279 GTGGCTGACCTCCCTCACTCAGCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATC 338
DB 361 GCGGACGACAAACCCCTTCTCGACCTCAGTGTCTACTTCTGCTGTGCTCGATATC 420
QY 339 C---ACAGCTGGAGTAGAAGCGGCGCTACTTTGCTGCACTGTGCTGTGCTGAGC 395
DB 421 CGAGCTGCCCTCAGTGTTCGCCAAGCGCGGTGCTGTGTACACTGTGCATGGGGTAAGC 480
QY 396 CGCTCAGCTGCCCTGTGCTGCTGCTACCTCATGAAGTACCACGCCATGCTCCCTGGAC 455
DB 481 CGCTTCCACACTGTCTCGGCTCTCTCATGTATCATGACATGACGCTGGTAGAG 540
QY 456 GCCCAGCTGGAGGACAGTATGCGGCCCATCATCGACCCCAACAGCGGCTTTTGGGAG 515
DB 541 GCCATCCAGAGGTGACGAGCCCGCAATATCTGCCTAACTCAGCGCTTCTCCGCGAG 600
QY 516 C 516
DB 601 C 601

RESULT 13

US-09-704-139-1

; Sequence 1, Application US/09704139

; Patent No. 6420153

; GENERAL INFORMATION:

; APPLICANT: Kapeller, Rosanna

; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR

; FILE REFERENCE: 10448-018001

; CURRENT APPLICATION NUMBER: US/09/704,139

; CURRENT FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: US 60/185,772

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 1390

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (329)..(859)

; NAME/KEY: misc_feature

; LOCATION: (1)..(1390)

; OTHER INFORMATION: n = a, t, g, or c

US-09-704-139-1

Query Match 6.3%; Score 52.2; DB 4; Length 1390;
Best Local Similarity 44.8%; Pred. No. 4.6e-06;

Matches 201; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 147 CTGTATATACAGCAATGTGTGGCGCAACAACAAAGCTCATGTCTAGCAACACAGATC 206
DB 404 CTCTTCTCGGAGCGCGGCGGCGGAGGAGGAGCTGGCGCGCGGGGAGTC 463
QY 207 ACCATGGTCATCAATGTCTCAGTGGAGTAGTGAACACCTTTGTATGAGGATATCCAGTAC 266
DB 464 ACCTGTGCGTCAAGCTCTCCCGCAGCAGCCCGCGCGCGCGCGCGGTGGCAGAG 523
QY 267 ATCAGGTACTGTGGCTGACTCCCTAACTCAGCTCTGTGACTTCTTTGACCTATT 326
DB 524 CTGCGGTGCGCGTGTTCGACGACCCCGGCTGAGGACCTGCTGGCGCACCTGGAGCCACG 583
QY 327 GCTGACCATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGCT 386
DB 584 TGGCGCCCATGGAGCGCGGCTGGCGCGCGCGCGCGCGCTGCTGCTACTTGAAGAAC 643
QY 387 GGTGTGAGCGCTCAGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
DB 644 GCGCGCAGCGCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 703
QY 447 CTGCTGAGCGCCACACAGCTGGACCAAGTATGCGCGCGCTGCTGCGCGCTGCTGCGCG 506
DB 704 CTGCGAAGCGCTTCCAGATGTTGAAGAGCGCTGCGCGCTGCTGCGCGCTGCTGCGCG 763
QY 507 TTTTGGGAGCAGCTCATCCACTATGATTTCCAAATTTTGGCAAGAACACTGTGCAATG 566
DB 764 TTTGCTGCTCAGCTGCTGAGAGTATGAGAGGCGCTTCCAGCGCGCGCTGCTGCTGCT 823
QY 567 GTCAGTTCCCGCAGTGAATGATCCCTGA 595
DB 824 GAGCCCCCAGCCTTAGGTTGGCGCTGA 852

RESULT 14
US-08-530-290-9
; Sequence 9, Application US/08530290
; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.

```

; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-071-035-183

Query Match 4.6%; Score 38.2; DB 4; Length 970;
Best Local Similarity 53.7%; Pred. No. 0.072;
Matches 79; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 579 ATTGTTGATCTTACACAAAGATCCAAACTTGAACATTCCTACTTTTGTGTATACAGAAAA 738
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 ATTGTCCTCTCTACTTCGGATGGAACTTGCCCAAGCTGTTTTTAAGTCTTCACTGTA 328
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 739 AACAGATGATGCCCTTTTATGAGCACAACAAAGAGTTCCTGTAGCTTTTTAACTTTATAATC 798
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 ATTTTAAAGTGGGCTTTTAGACCTGCATCATAGCTGCTCTCTTTTAACTGTTTCTTG 268
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 799 CATTTTTTTTCAGATTAAACTAATTGT 825
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 AATGTTTTTTTCAGTTAAACCGCTTGT 241
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: April 21, 2003, 17:55:20
Job time : 57 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 16:45:49 ; Search time 1070 Seconds
(without alignments)
12623.398 Million cell updates/sec

Title: US-09-527-376-1

Perfect score: 834

Sequence: 1 ctttctctgtattttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_othr.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	611	73.3	611	10	BE222374 hul1d05.x
C 2	543.6	65.2	592	9	AA527292 ng39g09.s
C 3	541.2	64.9	547	9	AA915932 on18c06.s
C 4	495.8	59.4	746	10	BE563213 601335506
C 5	490.2	58.8	495	9	AI215158 qp41h06.x
C 6	464.6	55.7	489	9	AA926744 om25a02.s

C	7	441	52.9	441	9	AI283262
	8	423.6	50.8	1161	11	AK015917
	9	413.8	49.6	446	13	BI037896
	10	359.8	43.1	894	14	BQ934977
	11	356.4	42.7	679	12	BG333140
C	12	341.8	41.0	457	10	AW972251
	13	324.6	38.9	744	13	BI907369
	14	313.4	37.6	840	11	AK007061
	15	305.2	36.6	784	13	BI914744
	16	302.6	36.3	625	10	BB661786
	17	291.8	35.0	429	12	BF559942
	18	290.2	34.8	674	12	BG423899
	19	286.8	34.4	813	12	BG619206
C	20	283.4	34.0	299	9	AI916043
	21	280.2	33.6	292	9	AA356476
	22	276.4	33.1	1102	13	BM550612
	23	269.2	32.3	563	10	AW844028
C	24	229	27.5	234	9	AI470481
C	25	226	27.1	562	10	AW843834
C	26	217.8	26.1	493	9	AI025489
C	27	201.4	24.1	491	9	AA813372
	28	197.4	23.7	925	12	BF160291
	29	191	22.9	818	13	BI464358
	30	190.6	22.9	571	9	AL711163
	31	187.2	22.4	1089	11	AK009744
	32	185.4	22.2	826	13	BI556609
C	33	184.2	22.1	460	9	AA725688
	34	183	21.9	1021	14	BQ951245
C	35	182.8	21.9	354	12	BF471540
	36	181	21.7	649	10	BB589466
	37	180.4	21.6	564	10	BE235660
	38	177.2	21.2	1016	9	AL527826
	39	175.4	21.0	876	12	BF130916
	40	175.2	21.0	1095	14	BM928368
	41	175	21.0	884	13	BI768742
C	42	173.8	20.8	188	10	AW372085
	43	173.4	20.8	949	14	BQ939586
	44	171.4	20.6	634	12	BG765121
C	45	171.2	20.5	465	9	AA813123

ALIGNMENTS

RESULT 1
BE222374/c
LOCUS hul1d05.xl NCI-CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3166281 3'
DEFINITION similar to TR:O95147 O95147 MKP-1 LIKE PROTEIN TYROSINE PHOSPHATASE
; mRNA sequence.
ACCESSION BE222374
VERSION BE222374
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 611)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco

611 bp mRNA linear EST 03-JUL-2000
hul1d05.xl NCI-CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3166281 3'
similar to TR:O95147 O95147 MKP-1 LIKE PROTEIN TYROSINE PHOSPHATASE
; mRNA sequence.

ACCESSION BE222374
VERSION BE222374
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 611)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco

```
FEATURES
source
High quality sequence stop: 485.
Location/Qualifiers
1. .611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3166281"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI-CGAP-Lu24 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT: 151 a 140 c 172 g 148 t
ORIGIN
Query Match 73.3%; Score 611; DB 10; Length 611;
Best Local Similarity 100.0%; Pred. No. 1.5e-158;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 CTCTCGCAGATAACCAAAAGCCTGTATATCAGCAATGTGTGCGCCGCAACAACAGCTC 185
Db 611 CTCTCGCAGATAACCAAAAGCCTGTATATCAGCAATGTGTGCGCCGCAACAACAGCTC 552
QY 186 ATGCTGTCTAGCAACACAGATCACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACC 245
Db 551 ATGCTGTCTAGCAACACAGATCACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACC 492
QY 246 TTGTATCAGGATATCCAGTACATGCAGGTACCTGTGTGCTGACCTCCCTCACTCAGCTCTC 305
Db 491 TTGTATCAGGATATCCAGTACATGCAGGTACCTGTGTGCTGACCTCCCTCACTCAGCTCTC 432
QY 306 TGTGACCTCTTTGACCCCTATTGTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGT 365
Db 431 TGTGACCTCTTTGACCCCTATTGTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGT 372
QY 366 ACTTTGCTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 425
Db 371 ACTTTGCTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 312
QY 426 ATGAAGTACACGCGCATGTCCCTGTGGAGCGCCACACAGCTGGACCAAGTTCATGCCGCC 485
Db 311 ATGAAGTACACGCGCATGTCCCTGTGGAGCGCCACACAGCTGGACCAAGTTCATGCCGCC 252
QY 486 ATCATCGACCCCAACAGCGCGCTTTTGGAGCAGCTCATCCACTATGAGTTTCCAAATTTGTT 545
Db 251 ATCATCGACCCCAACAGCGCGCTTTTGGAGCAGCTCATCCACTATGAGTTTCCAAATTTGTT 192
QY 546 GGCAAGAACACTGTGCACATGGTTCAGTTCCCTCAGTGGGAATGATCCCTGACATCTATGAG 605
Db 191 GGCAAGAACACTGTGCACATGGTTCAGTTCCCTCAGTGGGAATGATCCCTGACATCTATGAG 132
QY 606 AAGGAAGTCCGTTTGATGATTTCCACTGTGAGCCATCCACAGCGCCCTGATTTGGAGTCA 665
Db 131 AAGGAAGTCCGTTTGATGATTTCCACTGTGAGCCATCCACAGCGCCCTGATTTGGAGTCA 72
QY 666 GAGGTACAGATCTATTGTTGATCTTACCAAGATCCAAATTTGAACATTTCTACTTTTGT 725
Db 71 GAGGTACAGATCTATTGTTGATCTTACCAAGATCCAAATTTGAACATTTCTACTTTTGT 12
QY 726 TGATACAGAAA 736
Db 11 TGATACAGAAA 1
RESULT 2
AA527292/c 592 bp mRNA linear EST 05-AUG-1997
LOCUS
```

```

|||||
Db 293 TCCCGAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCGTTTGATGATTCACATG 234
QY 633 TGAGCCATCCCGAGGAGCCCTGCATGGAGTGCAGAGGTACAGATCTATTGTTGATCTTAC 692
Db 233 TGAGCCATCCCGAGGAGCCCTGCATGGAGTGCAGAGGTACAGATCTATTGTTGATCTTAC 174
QY 693 ACCAAGATCCAACTTGAACATCTACTTTTGTGATACAGAAAAAAGAGATGATGCCT 752
Db 173 ACCAAGATCCAACTTGAACATCTACTTTTGTGATACAGAAAAAAGAGATGATGCCT 114
QY 753 TTTATGAGCACAAAAAAGAGTGTCTGATCTTTTAACTTTTATAATCCATTTTTTTTTCAGA 812
Db 113 TTTATGAGCACAAAAAAGAGTGTCTGATCTTTTAACTTTTATAATCCATTTTTTTTTCAGA 54
QY 813 TTAACATAATTGTGAGATGGTG 834
Db 53 TTAACATAATTGTGAGATGGTG 32

RESULT 3
AA915932/c
LOCUS
DEFINITION
  AA915932 547 bp mRNA linear EST 23-JUN-1998
  on18c06.s1 NCI_CGAP_lu5 Homo sapiens cDNA clone IMAGE:1557034 3'
  similar to TR:Q91790 Q91790 MAP KINASE PHOSPHATASE ;, mRNA
  sequence.
ACCESSION
  AA915932
VERSION
  AA915932.1 GI:3055324
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  (bases 1 to 547)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/dbbrp/image/image.html
  Insert Length: 834 Std Error: 0.00
  Seq primer: -40ml3 fwd. Et from Amersham
  High quality sequence stop: 451.
  Location/Qualifiers
  1..547
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:1557034"
  /clone_lib="NCI_CGAP_lu5"
  /tissue_type="carcinoid"
  /lab_host="DH10B"
  /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
  modified polylinker; 1st strand cDNA was prepared from
  neuroendocrine lung carcinoid, and was then primed with a
  Not I - oligo(dT) primer. Double-stranded cDNA was ligated
  to Eco RI adaptors (Pharmacia), digested with Not I and
  cloned into the Not I and Eco RI sites of the modified
  pT7T3 vector. Library is normalized. Library was
  constructed by Bento Soares and M. Fatima Bonaldo. "
  BASE COUNT 142 a 126 c 153 g 126 t
  ORIGIN
  Query Match 64.9%; Score 541.2; DB 9; Length 547;
  Best Local Similarity 99.5%; Pred. No. 3.5e-139;
  Matches 543; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

|||||
QY 187 TCGTGTCTAGCAACAGATCACCATGTCATCAATGTCAGTGGAGTAGTGAACACCT 246
Db 546 TCGTGTCTAGCAACAGATCACCATGTCATCAATGTCAGTGGAGTAGTGAACACCT 487
QY 247 TGTATGAGGATATCCAGTACATGCAGGTACCTGTGGCTGACTCCCTAACTACAGTCTCT 306
Db 486 TGTATGAGGATATCCAGTACATGCAGGTACCTGTGGCTGACTCCCTAACTACAGTCTCT 427
QY 307 GTGACTTCTTTGACCCCTATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTA 356
Db 426 GTGACTTCTTTGACCCCTATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTA 367
QY 367 CTTTGTGTCACATGCTGCTGCTGAGCGGCTCAGCTGCCCTGTCCTGACCTACCTCA 426
Db 366 CTTTGTGTCACATGCTGCTGCTGAGCGGCTCAGCTGCCCTGTCCTGACCTACCTCA 307
QY 427 TGAAGTACCACGCCCATGTCCCTGCTGGAGCGCCACACAGTGGACCAAGTATGCGCGGCCA 486
Db 306 TGAAGTACCACGCCCATGTCCCTGCTGGAGCGCCATCAGTGGACCAAGTATGCGCGGCCA 247
QY 487 TCATCCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGATGTTCCAAATTTGTTG 546
Db 246 TCATCCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGATGTTCCAAATTTGTTG 187
QY 547 GCAAGAACACTGTGCACATGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGA 606
Db 186 GCAAGAACACTGTGCACATGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGA 127
QY 607 AGGAAGTCCGTTTGATGATTCCTGAGCCATCCACAGAGCCCTGCAATGGAGTCAG 666
Db 126 AGGAAGTCCGTTTGATGATTCCTGAGCCATCCACAGAGCCCTGCAATGGAGTCAG 67
QY 667 AGGTACAGATCTATTGTTGATCTTACCAAGATCCAAAGTTCGAATCTTACTTTGTT 726
Db 66 AGGTACAGATCTATTGTTGATCTTACCAAGATCCAAAGTTCGAATCTTACTTTGTT 7
QY 727 GATACA 732
Db 6 GATACA 1

RESULT 4
B5563213
LOCUS
DEFINITION
  B5563213 746 bp mRNA linear EST 15-AUG-2000
  601335306f1 NTH_MGC_39 Homo sapiens cDNA clone IMAGE:3689593 5',
  mRNA sequence.
ACCESSION
  B5563213
VERSION
  B5563213.1 GI:9807025
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  (bases 1 to 746)
REFERENCE
  NIH-MGC http://mgc.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLCM384 row: k column: 02
  High quality sequence stop: 700.
  Location/Qualifiers
  1..746
  /organism="Homo sapiens"
  /db_xref="taxon:9606"

FEATURES
  source

```

```

/clone="IMAGE:3689593"
/clone_lib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."
BASE COUNT 194 a 184 c 186 g 182 t
ORIGIN

Query Match 59.4%; Score 495.8; DB 10; Length 746;
Best Local Similarity 97.1%; Pred. No. 1.4e-126;
Matches 611; Conservative 0; Mismatches 7; Indels 11; Gaps 10;

Qy 207 ACCATGGTCATCATGTCTCAGTGGAGTAGTGAACACCTTTGTATGAGGATATCCAGTAC 266
Dy 1 ACCATGGTCATCAA-GTCTCAGTGGAGTAGTGAACACC-TGTATGAGGATATCCAGTAC 58

Qy 267 ATGAGGTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTCTTTGACCCCTATT 326
Dy 59 ATGAGGTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTCTTTGACCCCTATT 116

Qy 327 GCTGACCATATCCACAGC-GTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGC 385
Dy 117 GCTGACCATATCCACAGCGGTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGC 176

Qy 386 TGGTGTGAGCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCAACGCGCATGTC 445
Dy 177 TGGTGTGAGCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCAACGCGCATGTC 236

Qy 446 CTGCTGGACGCCACACGCTGGACCAAGTCATGCGGGGCCCATCATCGACCCCAACAGCGG 505
Dy 237 CTGCTGGACGCCACACGCTGGACCAAGTCATGCGGGGCCCATCATCGACCCCAACAGCGG 296

Qy 506 CTTTGTGGAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGACACTGTGCACAT 565
Dy 297 C-TTGTGGAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGACACTGTGCACAT 355

Qy 566 GGTGAGTTCCCAAGTGGGAATGATCCCTGACATCATGAGAAGGAAGTCCGTTTGTATGAT 625
Dy 356 GGTGAGTTCCCAAGTGGGAATGATCCCTGACATCATGAGAAGGAAGTCCG-TTGTATGAT 414

Qy 626 TCCACTGTGAGCCATCCACAGGCCCTGCAATGAGTTCAGAGTACAGATCTATTGTTG 685
Dy 415 TCCACTGTGAGCCATCCACAGGCCCTGCA-TGGAGTTCAGAGTACAGATCTA-TGTTG 472

Qy 686 ATCTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGTATACAGAAAAAACAGAT 745
Dy 473 ATCTTACACCAAGATCCAAACTTGAACATTTCTAC-TTGTGTATACAGAAAAAACAGAT 531

Qy 746 GATGCCCTTTATGAGCAAAAAAGAGTGTGTGAGCTTTTAACTTTTATATCCATTTT 805
Dy 532 GATGCCCTTTATGAGCAAAAAAGAG-TGCTGTAGCTTTTAACTTTTATATCCATTTT 590

Qy 806 TTTGAGATTAACATTAATGTGAGATGGTG 834
Dy 591 TTTGAGATTAACATTAATGTGAGATGGTG 619

RESULT 5
AI215158/c AI215158 495 bp mRNA linear EST 02-FEB-1999
LOCUS qp41ln06.x1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1925627 3'
DEFINITION similar to WP:FI3D11.3 CE04389 PROTEIN TYROSINE-PHOSPHATASE FAMILY
; mRNA sequence.
ACCESSION AI215158
VERSION AI215158.1 GI:3778759
KEYWORDS EST.

```

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 495)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1305 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 444.
Location/Qualifiers
1..495
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1925627"
/clone_lib="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 125 a 113 c 138 g 119 t
ORIGIN

Query Match 58.8%; Score 490.2; DB 9; Length 495;
Best Local Similarity 99.4%; Pred. No. 4.9e-125;
Matches 492; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 273 GTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTTCTTTGACCCCTATTGCTGAC 332
Dy 495 GTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTTCTTTGACCCCTATTGCTGAC 436

Qy 333 CATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGTGCTGCACTGTGCTGCTGGTGTG 392
Dy 435 CATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGTGCTGCACTGTGCTGCTGGTGTG 376

Qy 393 AGCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACACAGCCATGTCCTGCTG 452
Dy 375 AGCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACACAGCCATGTCCTGCTG 316

Qy 453 GACGCCACAGCTGGACCAAGTCATGCGGCCCATCATCCGACCCAAACAGCGGCTTTTGG 512
Dy 315 GACGCCACAGCTGGACCAAGTCATGCGGCCCATCATCCGACCCAAACAGCGGCTTTTGG 256

Qy 513 GAGCAGCTCATCCACTATGAGTTTCCAATTTTGGCAAGAACACTGTGCACATGCTCAGT 572
Dy 255 GAGCAGCTCATCCACTATGAGTTTCCAATTTTGGCAAGAACACTGTGCACATGCTCAGT 196

Qy 573 TCCCACTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGATGATTCACATG 632
Dy 195 TCCCACTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGATGATTCACATG 136

Qy 633 TGACCCATCCACAGCGCCCTGCAATGAGTACAGAGTCTATTTGTTGATCTTAC 692
Dy 135 TGACCCATCCACAGCGCCCTGCAATGAGTACAGAGTCTATTTGTTGATCTTAC 76

```

QY 693 ACCAAGTCCAACTTGAACATCTACTTTTGTGTATACAGAAAAACAGATGATGCCT 752
|||||
Db 75 ACCAAGATCCAACTTGAACATCTACTTTTGTGTATACAGAAAAACAGATGATGCCT 16
QY 753 TTTATGAGCACAAAA 767
|||||
Db 15 TTTATGAGCACAAAA 1

RESULT 6
AA926744/c
LOCUS
DEFINITION
  aa25a02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
  IMAGE:1542026 3' similar to WP:F26A3.4 CE09669 PROTEIN-TYROSINE
  PHOSPHATASE ;, mRNA sequence.
ACCESSION
  AA926744
VERSION
  AA926744.1 GI:3075641
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 489)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  This clone is available royalty-free through LNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1301 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 394.
  Location/Qualifiers
    1..489
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1542026"
    /clone_lib="Soares_NFL_T_GBC_S1"
    /lab_host="DH10B"
    /note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
    a modified polylinker; Site_1: Not I; Site_2: Eco RI;
    Equal amounts of plasmid DNA from three normalized
    libraries (fetal lung NbHL19w, testis NHT, and B-cell
    NCI-CGAP_GCB1) were mixed, and ss circles were made in
    vitro. Following HAP purification, this DNA was used as
    tracer in a subtractive hybridization reaction. The driver
    was PCR-amplified cDNAs from pools of 5,000 clones made
    from the same 3 libraries. The pools consisted of
    I.M.A.G.E. clones 297480-302087, 682632-687239,
    726408-728711, and 729096-731399. Subtraction by Bento
    Soares and M. Fatima Bonaldo."
  135 a 108 c 123 g 123 t

BASE COUNT
ORIGIN
  Query Match 55.7%; Score 464.6; DB 9; Length 489;
  Best Local Similarity 99.0%; Pred. No. 6.2e-118;
  Matches 478; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 352 TGAAGCAGGGCCGTACTTTGCTGCACGTGCTGCTGTGTGAGCCGCTCAGCTGCCTGT 411
|||||
Db 482 TTAAGCAGGGCCGTACTTTGCTGCACGTGCTGCTGTGTGAGCCGCTCAGCTGCCTGT 423
QY 412 GCCTCGCTACCTCANTAGTACCAGCCATGTCCTGCTGTGAGCCGCCACACGTGGACCA 471
|||||
Db 422 GCCTCGCTACCTCANTAGTACCAGCCATGTCCTGCTGTGAGCCGCCACACGTGGACCA 364
QY 472 AGTCATGCGCGCCCATCATCCGACCAACAGCGGCTTTTGGGAGCAGCTCATCCATNG 531
|||||
Db 363 AGTCATGCGCGCCCATCATCCGACCAACAGCGGCTTTTGGGAGCAGCTCATCCATNG 304
QY 532 AGTTCCAATTGTTGGCAAGAACACTGTGCACATGCTCAGTTCCTCCCGAGTGGGAATGATCC 591

```

```

Db 303 AGTTCCAATTGTTGGCAAGAACACTGTGCACATGCTCAGTTCCTCCCGAGTGGGAATGATCC 244
|||||
QY 592 CTGACATCTATGAGAAGGAGTCGCTTTGATGATTCACATGTCAGCATCCACGAGCCC 651
|||||
Db 243 CTGACATCTATGAGAAGGAGTCGCTTTGATGATTCACATGTCAGCATCCACGAGCCC 184
QY 652 CTGCATTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAA 711
|||||
Db 183 CTGCATTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAA 124
QY 712 CATCTACTCTTTGTTGATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGA 771
|||||
Db 123 CATCTACTCTTTGTTGATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGA 64
QY 772 GTTGTCTAGCTTTTAACTTTTATAATCCATTTTTTTTCAGATTAACATAATGATGATG 831
|||||
Db 63 GTTGTCTAGCTTTTAACTTTTATAATCCATTTTTTTTCAGATTAACATAATGATGATG 4
QY 832 GTG 834
|||||
Db 3 GTG 1

RESULT 7
AI283262/c
LOCUS
DEFINITION
  AI283262
  qk50g08.x1 NCI-CGAP_Co8 Homo sapiens cDNA clone IMAGE:1872446 3'
  similar to WP:F26A3.4 CE09669 PROTEIN-TYROSINE PHOSPHATASE ;, mRNA
  sequence.
ACCESSION
  AI283262
VERSION
  AI283262.1 GI:3921495
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 441)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1361 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 404.
  Location/Qualifiers
    1..441
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1872446"
    /clone_lib="NCI-CGAP_Co8"
    /tissue_type="adenocarcinoma"
    /lab_host="DH10B"
    /note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a
    modified polylinker; 1st strand cDNA was prepared from a
    colon adenocarcinoma, and was then primed with a Not I -
    oligo(dT) primer. Double-stranded cDNA was ligated to Eco
    RI adaptors (Pharmacia), digested with Not I and cloned
    into the Not I and Eco RI sites of the modified p7T73
    vector. Library is normalized. Library was constructed by
    Bento Soares and M. Fatima Bonaldo."
  109 a 103 c 122 g 107 t

BASE COUNT
ORIGIN

```



```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:4930527G07"
/db_xref="MGI:1908537"
/db_xref="taxon:10090"
/clone="4930527G07"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1..1161
/note="data source:SPTR, source key:Q9JLX7, evidence:ISS
similar to MAP KINASE PHOSPHATASE 6"
/db_xref="MGI:1922469"
BASE COUNT 279 a 324 c 295 g 263 t
ORIGIN

```

```

Query Match 50.8%; Score 423.6; DB 11; Length 1161;
Best Local Similarity 81.8%; Pred. NO. 1.6e-106;
Matches 489; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 61 CGGCCTTGATGACAGACACCTCGTGGCTTCCAGTTCAGTTCGGGAGCCCTCAGTCA 120
Db 324 CCACCTTGATGACTTCACCGTGGAGTCCCTCCAGTTCAGATCCCGCCAGCCCTCGATCA 383

QY 121 GGGGCTCTCGGAGATAACAAAGCCTGTATATACAGCAATGGTGGCGGCAACAACA 180
Db 384 GAGGCTCTCGGAGATACCAAGAGCCTGTTATCAGCAATGGTGGCTGGCAACAACA 443

QY 181 AGCTCATGCTGTCTAGCAACACAGATCACCATGGTTCATCATATGCTCAGTGGAGTAGTGA 240
Db 444 AGCTCTCATGCTGTCCAGCAATCAGATCAACACAGTCAACAGCTCTCAGTAGAGTAGCAA 503

QY 241 ACACCTTGATGAGGATATCCAGTACATGACAGTACCTGCTGGTACTCCCTTAACCTAC 300
Db 504 ACACCTTGATGAGGATATCCAGTATGTCAGAGTGGCTGTGGTTCGATGGCGGCTGCGAC 563

QY 301 GTCTGTGTGACTTCTTTGACCCCTATTGTCGACCATATCCACAGCGTGGAGATGAACGAG 360
Db 564 GCCTCTCAATTTCTTCGATTCCGTTGGCGGCGGATATCCATTCGTTGGTGGAGATGCAAGG 623

QY 361 GCGCTACTTTGTGCTACTGTGCTGTGTGTGAGCGGCTCAGCTGGCCCTGTGCTGCGCT 420
Db 624 GCGGACACATGTTGCAATGCTGTGCTGTGGGTGAGCGGCTCAGCTGGCTGTGCTGCGCT 683

QY 421 ACCTCATGAGTACACAGCCATGTCCTGCTGGAGCCGACAGCTGGAGGACCAAGTATGCC 480
Db 684 ACCTCATGAGTACCAATGTCATGTCCTGTTAGATGCCACACTTGGACCAAGTATGCA 743

QY 481 GGGCCATCATCGACCCACAGCGGCTTTGGGAGCAGCTCATCCACTATGAGTTCCAAAT 540
Db 744 GGGCCATCATCGGCCACACAGTGGCTTTGGGAGCAGCTCATCCATACGAGTTGCAGC 803

QY 541 TGTGTTGGCAAGAACACTGTGCATGGTTCAGTTCCCGAGTGGGAATGATCCCTGACATCT 600
Db 804 TCTTTGGCAAGAATAACAATGCAGATGATGACTCGCCTATGGGAGTGGAGATCCCGAGATCT 863

QY 601 ATGAGAGGAAGTCCGTTTGATGATTCACCTGTGAGCCATCCACAGAGCCCTGCATT 658
Db 864 ACAGAGGAGGAGAGCCGTTTGATGATCCCACTGTAAAGCCAGCCACCAAGTAACTCCAGT 921

```

```

RESULT 9
BI037896
LOCUS
DEFINITION CM4-NT0290-150101-681-a02 NT0290 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI037896
VERSION BI037896.1 GI:14444522
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS 1 (bases 1 to 446)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongseneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM4&t2=CM4-NT0290-
150101-681-a02&t3=2001-01-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 47
High quality sequence stop: 446.
FEATURES
Location/Qualifiers
1..446
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NT0290"
/dev_stage="Adult"
/note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 97 a 141 c 107 g 101 t
ORIGIN

```

```

Query Match 49.6%; Score 413.8; DB 13; Length 446;
Best Local Similarity 99.3%; Pred. No. 7.7e-104;
Matches 426; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 98 TCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACAAAGCCTGTATATCAG 157
Db 19 TCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAAAAT-ACCAAAGCCTGTATATCAG 77

QY 158 CAATGTTGGCGGCGCAACAAGCTCATGCTGTAGCAACCATCACCATGTCAT 217
Db 78 CAATGTTGGCGGCGCAACAAGCGCATGCTGTAGCAACCATCACCATGTCAT 137

QY 218 CAATGTTGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATGACGTACC 277
Db 138 CAATGTTGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATGACGTACC 197

QY 278 TGTGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTTGACCCCTATTGTCGACCATAT 337
Db 198 TGTGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTTGACCCCTATTGTCGACCATAT 257

QY 338 CCACAGCTGGAGATGAAGCAGGCGGTACTTTGTCGACTGTGCTGTGCTGTGAGCGG 397
Db 258 CCACAGCTGGAGATGAAGCAGGCGGTACTTTGTCGACTGTGCTGTGCTGTGAGCGG 317

QY 398 CTCAGTCCCTGTGCTCGCTACTCTCATGAAGTACCAAGCCATGCTCCCTGCTGACCC 457
Db 318 CTCAGTCCCTGTGCTCGCTACTCTCATGAAGTACCAAGCCATGCTCCCTGCTGACCC 377

QY 458 CCACAGCTGGAGCAAGTATATGCGGCGCCATCATCCGACCAACAGCGGCTTTTGGAGCA 517
Db 378 CCACAGCTGGAGCAAGTATATGCGGCGCCATCATCCGACCAACAGCGGCTTTTGGAGCA 437

```

```

QY 518 GCTCATCCA 526
|||||
Db 438 GCTCATCCA 446

RESULT 10
BO934977      894 bp      mRNA      linear      EST 21-AUG-2002
LOCUS      AGENCOURT_8881260 NCI_CGAP_Co24 Mus musculus cDNA clone
DEFINITION      IMAGE:6396450 5', mRNA sequence.
ACCESSION      BO934977
VERSION      BO934977.1 GI:22350360
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 894)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13892 row: 1 column: 19
High quality sequence stop: 676.
FEATURES      Location/Qualifiers
source      1..894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6396450"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      229 a 229 c 220 g 216 t
ORIGIN

Query Match      43.1%; Score 359.8; DB 14; Length 894;
Best Local Similarity 81.9%; Pred. No. 8e-89;
Matches 415; Conservative .0; Mismatches 92; Indels 0; Gaps 0;

QY 152 TATCAGCAATGGTGTGCGCCGACACACACAGCTCATGCTGTCTAGCAACACGATCACCAT 211
|||||
Db 2 TATCAGCAATGGTGTGCTGCCAACACACAGCTCTCTACTGTCTCCAGCAATCAGATCACCAC 61

QY 212 GGTATCAATGCTCTCAGTGGAGTAGTGAACACCTCTGTATGAGGATATCAGTACATGCA 271
|||||
Db 62 AGTCATCAACGCTCTCAGTAGNGTAGCAACACCTCTTACGAGGATATCAGTATGCA 121

QY 272 GGTACCTGTGGCTGACTCCCTCACTACCTACCTGCTGTGACTTCTTTGACCCCTATTTGCTGA 331
|||||
Db 122 GGTGCTGTGTGCTGATCGCCGCGTCGACCGCTCTCCAATTTCTTCGATTCGGTTCGCGA 181

QY 332 CCATATCCACAGCTGTGAGATGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
|||||
Db 182 CCGTATCCATTTGCTGTGAGATGAGAGAGGCGGCACACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 241

QY 392 GAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
|||||
Db 242 GAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

QY 452 GGACGCGCCACAGCTGGACCAAGTATCATGCCGCGCCCATCATCCGACCCCAACAGCGGCTTTTG 511
|||||

```

```

Db 302 AGATGCCACACTTTGGACCAAGTCATCGAGCCCATCATCGGCCCAACAGTGGCTTTTG 361
|||||
QY 512 GGAGCAGCTCATCCACTATGAGTTCCTCAATGTTTGGGAAGACACTGTGCACATGGTCAG 571
|||||
Db 362 GGAGCAGCTCATCCATTTACGAGTTCGAGCTCTTTTGGCAAGAATACAATGCAGATGATGA 421
|||||
QY 572 TTCCCGAGTGGGATGATCCCTGACATCTATGAGAAGGAAGTCGTTTGTGATGATTCCTACT 631
|||||
Db 422 CTCGCCAATGGAGCGATCCAGACATCTAGGAGAAGAGACCCGTTTGTGATGCCACT 481
|||||
QY 632 GTGAGCCATCCACGAGCCCTCTGCAAT 658
|||||
Db 482 GTAAGCCAGCCACCACTTAATCCAGT 508

RESULT 11
BG333140      679 bp      mRNA      linear      EST 27-FEB-2001
LOCUS      602430837F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4548585 5',
DEFINITION      mRNA sequence.
ACCESSION      BG333140
VERSION      BG333140.1 GI:13139578
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 679)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CMI238 row: j column: 10
High quality sequence stop: 679.
FEATURES      Location/Qualifiers
source      1..679
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4548585"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      145 a 195 c 187 g 151 t 1 others
ORIGIN

Query Match      42.7%; Score 356.4; DB 12; Length 679;
Best Local Similarity 96.8%; Pred. No. 6.8e-88;
Matches 363; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 23 CATTTTGGTGTTCGCTGACTGCTGACCACTGACCCACCCCTTGTGATGACAGCACCTC 82
|||||
Db 276 CAGCCTTGGTGTTCGCTGACTGCTGACCACTGACCCACCCCTTGTGATGACAGCACCTC 335

QY 83 GTGTGCTTCCCACTTCCAGTTCGGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 142
|||||
Db 336 GTGTGCTTCCCACTTCCAGTTCGGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 395
|||||

```


[illegible][illegible]

```

FEATURES
  High quality sequence stop: 743.
  Location/Qualifiers
    source
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="IMAGE51253"
      /clone_lib="NH MGC118"
      /tissue_type="leukocyte"
      /lab_pos="ps103"

```

```

/notes=vector; pcwm-sp0k10; site_1: NotI; site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

BASE COUNT      152 a      222 c      216 g      154 t
ORIGIN

Query Match      38.9%; Score 324.6; DB 13; Length 744;
Best Local Similarity 98.88; Pred. No. 4.6e-79;
Matches 327; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CATTCCTGGTGTTCGTGACTGCTGACCACTGACCACCGCCCTTGATGACAGCACCCCTC 82

```



```
/db_xref="GI:12840422"
/db_xref="MGD:GI:1920797"
/translators="MTTASCFPSOATQDNIYCLSOITASLFSNSAVANDKLTLSN
NHITTLINSAEVNFFEDIQVQVPVSDAPNSLYLDFEDPTADHGHVEMNGRFL
LHCAAGVRSATLCLAYLMYKHNHTLDAHTWTCTRPRIIPNPNNGFEQLIHYEFKLF
SRNTRVMIYSPILNPIYKEAYLMEIM"
polya_signal      820..825
                  /note="putative"
polya_site        840
                  /note="putative"
BASE COUNT      241 a 206 c 175 g 218 t
ORIGIN
```

```
Query Match      37.6%; Score 313.4; DB 11; Length 840;
Best Local Similarity 70.7%; Pred. No. 6e-76;
Matches 432; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

Qy 53 CTGACCCACCCCTTGATGACGACGACCCCTGCTGTGCTCCCTCCAGTTCA---GTTCCGGCA 109
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 77 CTGACCTGTACCGGTGATGACACAGCATCTTGTATCTTTCCGCTCAAGCTACCAACA 136
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 110 GCCCTCAGTCAGCGCTCTCGAGATACCAAAAGCCTGTATATCAGCAATGGTGTGGC 169
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 137 GGATAACATCTATGCTCTCTCAATAAATGCGCAGTCTTTTATCAGTAATCTGCGGT 196
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 170 CGCAACAACAAGCTCATGCTGTCTAGCAACGATACCATGCTCATCAATGCTCAGT 229
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 197 AGCTAACGACAAATCACCCTGTCTCAACACCATATCACCAGATCATCAATGCTCTGC 256
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 230 GGAGTAGTAGAACCTTTGTATGAGGATPATCCAGTACATGAGGTACCTGTGGTGA 289
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 257 AGAAGTAGTAATAGTCTTTTGAAGACATTCAGTACGTGCAAGTGCCTGAGCGATGC 316
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 290 CCCTAACCTCAGCTCTGTGACTCTTTGACCCATATGCTGACCATATCCACAGCGTGA 349
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 317 TCCTAAATCTTACCTCTATGACTTTTTCGACCCCATGCTGATCATCCACGCTGTGA 376
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 350 GATGAAGCAGGCGCTACTTTGTGCACTGTGCTGCTGTGAGCGCTCAGCTGCCT 409
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 377 AATGAGAATGGCGCAGCTGTGCACTGTGCGCAGAGATGAGCCGCTCAGCAACTCT 436
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 410 GTGCTCGCCCTACCTCATGAATACCAAGCCATGCTCCCTGCTGGACGCCACACGTGGAC 469
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 437 CTGCTTGCCTATCTCATGAATACCAATATGACCCCTGCTGGATGCACACATGGAC 496
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 470 CAAGTCATGCGGCCCATCATCCGACCAACAGCGGCTTTTGGGAGAGCTCATCCACTA 529
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 497 CAAGCTGCGCGCCCATCATCCGTCGCCAACAATGGCTTTTGGGAACAACCTGATTCATTA 556
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 530 TGAGTTCCAATTTGTTGGCAAGAACACTGTGCACATGCTCAGTTCCCGAGTGGGAATGAT 589
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 557 CGAGTTCAAGCTGTTTAGTAGGAATACCGTTCCGATGATCTCTCTATAGTCTGAT 616
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 590 CCCTGACATCATGAGAAGAGTCCGTTTGATGATNTCCATGTGAGCCATCCACAGAGC 649
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 617 TCCTAACATCTATGAAGAGGCGCTATTTAATGAGCTAATGTAATCCACCCATCTCT 676
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 650 CCCTGCATTGG 660
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 677 CCCTGATTGG 687
    |||||  ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
```

```
RESULT 15
LOCUS      BI914744
DEFINITION 603184117f1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248013 5',
            mRNA sequence.
ACCESSION  BI914744
VERSION    BI914744.1 GI:16195028
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 784)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1626 row: a column: 06
High quality sequence stop: 742.

FEATURES
source
Location/Qualifiers
1..784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5248013"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

BASE COUNT 164 a 226 c 232 g 162 t
ORIGIN

Query Match 36.6%; Score 305.2; DB 13; Length 784;
Best Local Similarity 96.9%; Pred. No. 1.e-73;
Matches 343; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
Qy 23 CATCTTGGTGTTCGCTGACTGTGACACTGACCCACCGCTTGTGATGACAGCACCTTC 82
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 408 CAGCCTTCGTGCTTCGCTGACTGTGCTGACACTGACCCACCGCTTGTGATGACAGCACCTTC 467
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 83 GTGTGCTTCCAGTTCAGTTCCTCCGACGCTTCAGTCCAGGCTTCAGTCCAGGCTTC 142
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 468 GTGTGCTTCCAGTTCAGTTCCTCCGACGCTTCAGTCCAGGCTTCAGTCCAGGCTTC 527
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 143 AAGCCTGTATATCAGCAATGTGTGCGCGCCCAACAAAGCTCATGCTGTCTAGCAACCA 202
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 528 AAGCCTGTATATCAGCAATGTGTGCGCGCCCAACAAAGCTCATGCTGTCTAGCAACCA 587
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 203 GATCACATGGTCAATGCTCAGTGAGGTAGTGAACCTTGTATGAGGATATCCA 262
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 588 GATCACATGGTCAATGCTCAGTGAGGTAGTGAACCTTGTATGAGGATATCCA 647
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 263 GTACATGCA-GGTACCTGTGGCTGACTCCCTTAACCTACAGT-CTCTGTGACTTCTTTGAC 320
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 648 GTACATGAGGCTGACTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGAG 707
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Qy 321 CCTATT-GCTGACCATATCCACAGGTGAGATGAAGCAGGCGCGCTACTTTGCT 373
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 708 CCTATTGGCTGACCATATCAACAGGTGAGATGAAGCAGGCGCGCTACTTACCT 761
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Search completed: April 21, 2003, 17:54:14
Job time : 1078 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:45:03 ; Search time 1834 Seconds
(without alignments)
13234.318 Million cell updates/sec

Title: US-09-527-376-1
Perfect score: 834
Sequence: 1 ctttctgtatttttgc.....aaactaattgtgatgtgtg 834

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.btgo.hum.*
- 40: em.btgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	805.6	96.6	1407	6	AX202236	Sequence
2	805.6	96.6	2454	9	AK056074	Homo sapi
3	804	96.4	39401	9	AC003072	Human BAC
4	704.6	84.5	775	6	AX086024	Sequence
5	647	77.6	1865	9	BC004110	Homo sapi
6	640	76.7	1335	9	BC030987	Homo sapi
7	575.6	69.0	2921	9	BC028724	Homo sapi
8	474.4	56.9	191511	2	AC125764	Rattus no
9	445	53.4	1655	9	AK025036	Homo sapi
10	423.6	50.8	2681	10	BC020036	Mus muscu
11	423.6	50.8	186768	2	AL807241	Mus muscu
12	423.6	50.8	205509	10	AL731853	Mouse DNA
13	369.8	44.3	908	6	AX086022	Sequence
14	362.6	43.5	140717	9	AL133545	Human DNA
15	362.4	43.5	935	6	AX040320	Sequence
16	343.4	41.2	889	9	AF143321	Homo sapi
17	343.4	41.2	889	9	AY121807	Homo sapi
18	322.6	38.7	177327	2	AC113681	Homo sapi
19	322.6	38.7	189620	2	AC126908	Rattus no
20	313.4	37.6	161816	10	AL773547	Mouse DNA
21	187.2	22.4	1425	10	BC002130	Mus muscu
22	187.2	22.4	196812	10	AL645615	Mouse DNA
23	187.2	22.4	201430	2	AL596447	Mus muscu
24	185.6	22.3	597	10	AF120113	Mus muscu
25	177.6	21.3	1167	9	AF120032	Homo sapi
26	177.6	21.3	1445	9	BC004448	Homo sapi
27	177.6	21.3	1454	9	BC000370	Homo sapi
28	177.6	21.3	1471	9	AF038844	Homo sapi
29	177.6	21.3	1474	9	BC001894	Homo sapi
30	177.6	21.3	1729	6	AR122778	Sequence
31	177.6	21.3	136222	9	AC004099	Homo sapi
32	151	18.1	461	11	G28106	human STS
33	137.6	16.5	2259	9	AK027210	Homo sapi
34	98.2	11.8	142093	9	AC026884	Homo sapi
35	98.2	11.8	170934	2	AL354684	Homo sapi
36	90	10.8	1059	5	AF026522	Gallus ga
37	88.8	10.6	1933	6	AX306137	Sequence
38	88.8	10.6	1933	10	MM3CH134M	Mouse mRNA
39	88.8	10.6	1941	10	BC006967	Mus muscu
40	88.2	10.6	1467	5	AF167296	Gallus ga
41	88	10.6	2000	6	AX337238	Sequence
42	88	10.6	2000	6	AX411130	Sequence
43	88	10.6	2000	9	HSCL100	H.sapiens C
44	88	10.6	2026	9	BC022463	Homo sapi
45	87	10.4	1209	6	AX369208	Sequence

ALIGNMENTS

RESULT 1	AX202236	AX202236	1407 bp	DNA	linear	PAT 30-AUG-2001
LOCUS	AX202236	Sequence 15 from Patent WO0153469.				
DEFINITION	AX202236	Sequence 15 from Patent WO0153469.				
ACCESSION	AX202236	AX202236.1	GI:15392011			
VERSION	AX202236.1	GI:15392011				
KEYWORDS		human.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		1 (bases 1 to 1407)				
		Bandman,O., Tang,Y.T., Azimzai,Y., Yue,H., Baughn,M.R.,				
		Hillman,J.L., Lal,P., Wang,E., Gandhi,A.R., Policky,J.L. and				
		Mathur,P.				

RESULT 2

```
QY 263 GTACATGAGGTTACCTGTGGGTGACTCCCTAACTACAGTCTCTGTGACTCTTTTGACCC 322
Db 696 GTACATGAGGTTACCTGTGGGTGACTCCCTAACTACAGTCTCTGTGACTCTTTTGACCC 755
QY 323 TATTGCTCACCATACACAGCGTGGAGATGAAGCAGGCGGTACTTTGCTGCACGTGTC 382
Db 756 TATTGCTCACCATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGCTGCACGTGTC 815
QY 383 TGCTGCTGTGAGCCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCGAGCCCAT 442
Db 816 TGCTGCTGTGAGCCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCGAGCCCAT 875
QY 443 GTCCCTGCTGGAGCCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCGAGCCCAT 502
Db 876 GTCCCTGCTGGAGCCGCTCAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCGAGCCCAT 935
QY 503 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCGAATTTGGCAAGACACTGTGCA 562
Db 936 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCGAATTTGGCAAGACACTGTGCA 995
QY 563 CATGCTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 622
Db 996 CATGCTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 1055
QY 623 GATTCACCTGTGAGCCATCCCGAGCCGCTGCAATGAGTGACAGGTACAGATCTATTG 682
Db 1056 GATTCACCTGTGAGCCATCCCGAGCCGCTGCAATGAGTGACAGGTACAGATCTATTG 1115
QY 683 TTGATCTTACACCAAGATCCAACTTGAACATCTACTTTTGTGATACAGAAAAACA 742
Db 1116 TTGATCTTACACCAAGATCCAACTTGAACATCTACTTTTGTGATACAGAAAAACA 1175
QY 743 GATGATGCTTTTATGACACAAAAAGATTGCTGTAGCTTTTAACTTTATATATCCATT 802
Db 1176 GATGATGCTTTTATGACACAAAAAGATTGCTGTAGCTTTTAACTTTATATATCCATT 1235
QY 803 TTTTTCAGATTAATACTAAATTTGTGAGATGGTG 834
Db 1236 TTTTTCAGATTAATACTAAATTTGTGAGATGGTG 1267

RESULT 3
AC003072/c 39401 bp DNA linear PRI 04-FEB-2000
LOCUS Human BAC clone CTA-963H5 from 22q12.1-qter, complete sequence.
DEFINITION AC003072
ACCESSION AC003072
VERSION AC003072.1 GI:2588642
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 39401)
AUTHORS Murray,J., Wohlmann,P., Bauer,C. and Courtney,L.
TITLE The sequence of H. sapiens BAC clone CTA-963H5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1997) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 3 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
```

Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_BR963H05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre chromosome 22 mapping group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22/>

SOURCE INFORMATION:

This clone is from the human BAC library described by U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBelBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone being sequenced to the left is RP1-56J10. The actual start of this clone is at base position 1 of CTA-963H5. This clone is part of an unanchored island, orientation is unknown.

FEATURES	Source
repeat_region	1. 39401 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="22" /map="22q12.1-qter" /clone="CTA-963H5" /clone_lib="BKB" complement(1142..1435) /rpt_family="ALU" complement(1750..2045) /rpt_family="ALU" 2937..3203 /rpt_family="ALU" complement(3758..4038) /rpt_family="ALU" 4595..4766 /rpt_family="ALU" 4915..5230 /rpt_family="ALU" 5239..5363 /rpt_family="ALU" 5385..5675 /rpt_family="ALU" 5692..5875 /rpt_family="ALU" 6019..6148 /rpt_family="ALU" 6625..6895 /rpt_family="ALU" complement(7143..7413) /rpt_family="ALU" complement(8074..8235) /rpt_family="ALU" complement(8610..8902) /rpt_family="ALU" complement(9218..9360)


```

|||||
Db 310 CATCCGACCAACAGCGGCTTTGGGAGCAGCTCATCCATGATGAGTTCCAAATTTGTTGG 369
QY 548 CAAGAACACATGTCACATGTCAGTCCCTCCAGTGGGAATGATCCCTGACATCTATGAGAA 607
Db 370 CAGAACACATGTCACATGTCAGTCCCTCCAGTGGGAATGATCCCTGACATCTATGAGAA 429
QY 608 GGAAGTCGCTTTGATGATTCACATGTCAGGATCCACAGAGCCCTGCAATGGAGTCAGA 667
Db 430 GGAAGTCGCTTTGATGATTCACATGTCAGGATCCACAGAGCCCTGCAATGGAGTCAGA 489
QY 668 GGTACAGATCTATGATGATCTTACACCAAGATCCAAACTTGAACTTACATCTTGTGTTG 727
Db 490 GGTACAGATCTATGATGATCTTACACCAAGATCCAAACTTGAACTTACATCTTGTGTTG 549
QY 728 ATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTA 787
Db 550 ATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTA 609
QY 788 ACTTTATAATCCATTTTTTTTCAGATTAATACTAATTTGATGATGGTG 834
Db 610 ACTTTATAATCCATTTTTTTTCAGATTAATACTAATTTGATGATGGTG 656

```

```

RESULT 6
BC030987
LOCUS
DEFINITION
Homo sapiens, similar to RIKEN CDNA 4930527G07 gene, clone
MGC:32658 IMAGE:4730282, mRNA, complete cds.
ACCESSION
BC030987
VERSION
BC030987.1 GI:21410213
KEYWORDS
MGC.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT
Contact: MGC help desk
Email: gcgaps@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 41 Row: m Column: 12
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction, Similarity but not identity to protein.

```

FEATURES
source
1. .1335
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:32658 IMAGE:4730282"
/tissue_type="placenta"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
460. .1026
CDS

```

```

/codon_start=1
/product="Similar to RIKEN CDNA 4930527G07 gene"
/protein_id="AAH30987.1"
/db_xref="GI:21410214"
/translation="MTAPSCAFVQFQRPQSVSGLSQITKSLYISNGVAANKMLSSN
QITMWINVSEVVNTLIEDIQMQVPVADSPNSRLCDFEDPIADHLSHVMKQGRLL
HCAAGVSRSAALCLAYLKMRYHAMSLLDAHTWTKSRPIIRPNSGFEQLIHVEFQLFG
KNTVMVSPVGMIPDIYEKEVRLMPL"
BASE COUNT 304 a 399 c 346 g 286 t
ORIGIN

Query Match 76.7%; Score 640; DB 9; Length 1335;
Best Local Similarity 98.5%; Pred. No. 2.6e-164;
Matches 646; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 23 CATTCTTGGTGTTCCTGACTGCTGACCACTGACCCACCGCTTGTGATGACAGCACCCCTC 82
Db 414 CAGCCTTCGTGCTTCGCTGACTGCTGACCACTGACCCACCGCTTGTGATGACAGCACCCCTC 473
QY 83 GTGTGCTTTCCCAAGTTTCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 142
Db 474 GTGTGCTTTCCCAAGTTTCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 533
QY 143 AAGCCTGTATATCAGCAATGGTGTGGCGGCCAACACAAAGCTCATGCTGTCTAGCAACCA 202
Db 534 AAGCCTGTATATCAGCAATGGTGTGGCGGCCAACACAAAGCTCATGCTGTCTAGCAACCA 593
QY 203 GATCACCATTGGTTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGATATCCA 262
Db 594 GATCACCATTGGTTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGATATCCA 653
QY 263 GTACATGCAGGTACCTGTGGCTGACTCCCTAACTACGCTCTCTGTGACTTCTTTGACCC 322
Db 654 GTACATGCAGGTACCTGTGGCTGACTCCCTAACTACGCTCTCTGTGACTTCTTTGACCC 713
QY 323 TATTGTGTACCATPATCCACAGCGTGGAGATGAAGCAGGCGGCTGACTTTGTGCTGACTGTGC 382
Db 714 TATTGTGTACCATPATCCACAGCGTGGAGATGAAGCAGGCGGCTGACTTTGTGCTGACTGTGC 773
QY 383 TGTGTGTGTGAGCGGCTCAGCTGCGCTGTGCTCGCTACCTCATGAAGTACCAACGCCAT 442
Db 774 TGTGTGTGTGAGCGGCTCAGCTGCGCTGTGCTCGCTACCTCATGAAGTACCAACGCCAT 833
QY 443 GTCCTGTGTGAGCGGCTCAGCTGCGGCTCATGCGGCGCCATCATCGAGCCCAACAG 502
Db 834 GTCCTGTGTGAGCGGCTCAGCTGCGGCTCATGCGGCGCCATCATCGAGCCCAACAG 893
QY 503 CGGCTTTTGGGAGCAGCTCATCCACTATGATGTTCCAAATGTTTGGCAAGAACACTGTGCA 562
Db 894 CGGCTTTTGGGAGCAGCTCATCCACTATGATGTTCCAAATGTTTGGCAAGAACACTGTGCA 953
QY 563 CATGGTCAGTTCCCGCAGTGGGAATGATCCCTGACATCTATGAGAAGAACTCGTTTGTAT 622
Db 954 CATGGTCAGTTCCCGCAGTGGGAATGATCCCTGACATCTATGAGAAGAACTCGTTTGTAT 1013
QY 623 GATTCCACTGTGAGCGCATCCACAGAGCCCTGCAATTTGGAGTCAGAGTACAGATCT 678
Db 1014 GATTCCACTGTGAGCGCATCCACAGAGCCCTGCAATTTGGAGTCAGAGCGCCCTG 1069

RESULT 7
BC028724
LOCUS
DEFINITION
Homo sapiens, similar to Unknown (protein for IMAGE:3689593), clone
MGC:33231 IMAGE:5269586, mRNA, complete cds.
ACCESSION
BC028724
VERSION
BC028724.1 GI:20381112
KEYWORDS
MGC.
SOURCE
Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2921)

```



```

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunou,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pichens,R., Primus,E., Fu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 191511)
Worley,K.C.
Direct Submission
Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 191511)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 2, 2002 this sequence version replaced gi:21630358.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: G2FB
Center clone name: CH230-229p3
-----
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143602 bases at least Q40
Consensus quality: 151894 bases at least Q30
Consensus quality: 158330 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1316: contig of 1316 bp in length
* 1317: 1416: gap of unknown length
* 1417: 3055: contig of 1639 bp in length
* 3056: 3155: gap of unknown length
* 3156: 4162: contig of 1007 bp in length
* 4163: 4262: gap of unknown length
* 4263: 5639: contig of 1377 bp in length
* 5640: 5739: gap of unknown length
* 5740: 6782: contig of 1043 bp in length
* 6783: 6882: gap of unknown length
* 6883: 8022: contig of 1140 bp in length
* 8023: 8122: gap of unknown length
* 8123: 9761: contig of 1639 bp in length
* 9762: 9861: gap of unknown length
* 9862: 11958: contig of 2097 bp in length
* 11959: 12038: gap of unknown length
*
* 12059: 13754: contig of 1696 bp in length
* 13755: 13854: gap of unknown length
* 13855: 15390: contig of 1536 bp in length
* 15391: 15490: gap of unknown length
* 15491: 16642: contig of 1152 bp in length
* 16643: 16742: gap of unknown length
* 16743: 18412: contig of 1670 bp in length
* 18413: 18512: gap of unknown length
* 18513: 20412: contig of 1900 bp in length
* 20413: 20512: gap of unknown length
* 20513: 22041: contig of 1529 bp in length
* 22042: 22141: gap of unknown length
* 22142: 23649: contig of 1508 bp in length
* 23650: 23749: gap of unknown length
* 23750: 25360: contig of 1611 bp in length
* 25361: 25460: gap of unknown length
* 25461: 27686: contig of 2226 bp in length
* 27687: 27786: gap of unknown length
* 27787: 29677: contig of 1891 bp in length
* 29678: 29777: gap of unknown length
* 29778: 32105: contig of 2328 bp in length
* 32106: 34956: contig of 2751 bp in length
* 34957: 35056: gap of unknown length
* 35057: 37267: contig of 2211 bp in length
* 37268: 37367: gap of unknown length
* 37368: 39228: contig of 1861 bp in length
* 39229: 39328: gap of unknown length
* 39329: 40783: contig of 1455 bp in length
* 40784: 40883: gap of unknown length
* 40884: 42941: contig of 2058 bp in length
* 42942: 43041: gap of unknown length
* 43042: 44879: contig of 1838 bp in length
* 44880: 44879: gap of unknown length
* 44880: 46226: contig of 1247 bp in length
* 46227: 46326: gap of unknown length
* 46327: 48011: contig of 1685 bp in length
* 48012: 48111: gap of unknown length
* 48112: 49971: contig of 1860 bp in length
* 49972: 50071: gap of unknown length
* 50072: 51849: contig of 1778 bp in length
* 51850: 51949: gap of unknown length
* 51950: 54291: contig of 2342 bp in length
* 54292: 54391: gap of unknown length
* 54392: 57401: contig of 3010 bp in length
* 57402: 57501: gap of unknown length
* 57502: 59171: contig of 1670 bp in length
* 59172: 59271: gap of unknown length
* 59272: 62524: contig of 3253 bp in length
* 62525: 62624: gap of unknown length
* 62625: 64053: contig of 1429 bp in length
* 64054: 64153: gap of unknown length
* 64154: 66493: contig of 2340 bp in length
* 66494: 66593: gap of unknown length
* 66594: 68444: contig of 1851 bp in length
* 68445: 68544: gap of unknown length
* 68545: 71479: contig of 2935 bp in length
* 71480: 71579: gap of unknown length
* 71580: 74791: contig of 3212 bp in length
* 74792: 74891: gap of unknown length
* 74892: 78202: contig of 3311 bp in length
* 78203: 78302: gap of unknown length
* 78303: 82035: contig of 3733 bp in length
* 82036: 82135: gap of unknown length
* 82136: 85344: contig of 3209 bp in length
* 85345: 85444: gap of unknown length
* 85445: 88700: contig of 3256 bp in length
* 88701: 88800: gap of unknown length
* 88801: 93292: contig of 4492 bp in length
* 93293: 93392: gap of unknown length
* 93393: 98296: contig of 4904 bp in length
* 98297: 98396: gap of unknown length
* 98397: 101597: contig of 3201 bp in length

```


Db 361 CCTTTTATGACACAAAAGAGTGTCTGTAGCTTTTAACTTTTATATCACTTTTTC 420

QY 810 AGATTAACCTAATGTGAGATGGTG 834

Db 421 AGATTAACCTAATGTGAGATGGTG 445

RESULT 10

BC020036 2681 bp mRNA linear ROD 07-AUG-2002

LOCUS Mus musculus, clone MGC:28218 IMAGE:3990751, mRNA, complete cds.

DEFINITION BC020036

ACCESSION BC020036.1 GI:18043292

VERSION MGC.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2681)

REFERENCE Strausberg, R.

AUTHORS Direct Submission

TITLE Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

JOURNAL NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK Contact: MGC help desk

COMMENT Email: crabbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 36 Row: 1 Column: 5

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..2681

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="CZECH II"

/clone="MGC:28218 IMAGE:3990751"

/tissue_type="Mammary tumor metastatized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."

/clone_lib="NCI_CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

342..908

/codon_start=1

/product="Unknown (protein for MGC:28218)"

/protein_id="AAH20036.1"

/db_xref="GI:18043293"

/translation="MTSPWSPFVQIPQPSIRGLSOTKSLFTSNGVAANNKLLSSN QITVLNVSEVANTVEDIQYQVPVDPAPVARLSNFFDSVADRIHVSVMOKGRLL HCAAGYSRSALCLAYLMKYHANSLVDAHTWTKSCRIPRNSGFWEQLIHVELQLFG KNTQMDSMPMGRIPDIYEKTRIMPL"

BASE COUNT 666 a 689 c 659 g 667 t

ORIGIN

Query Match 50.88; Score 423.6; DB 10; Length 2681;

Best Local Similarity 81.88; Pred. No. 6.6e-105;

Matches 489; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 61 CCGCTTTGATGACAGCACCCCTCGTGCCTTCCAGTTCCAGTTCGGCAGCCCTCAGTCA 120

Db 334 CCACCTTTGATGACTTCACCGTGGAGTGCCTTCCAGTTCCAGTTCAGATCCCGCAGCTTCGATCA 393

QY 121 GCGGCTCTCCAGATAACCAAAAGCCTGTATATACACAATGGTGTGGCGCCCAACAACA 180

Db 394 GAGGCTCTCCAGATCACCAAGAGCCTGTTTATCAGCAATGGTGTGGTCCCAACAACA 453

QY 181 AGCTCATGCTGTCTAGCAACAGATCACCATGCTCATCAATGCTCAGTGGAGTAGTGA 240

Db 454 AGCTCTTACTGTCCAGCAATCAGATCACCAAGTTCATCAACGCTCTAGTAGAGTAGCAA 513

QY 241 ACCTTTGATGAGGATATCCAGTACATCCAGTAGTACCTGTGGCTGACTCCCTCAATCCAC 300

Db 514 ACACCTTCTAGGAGATATCCAGTATGTGCAGGTGCCTGTGGTGCATGCCCGCTCGCAC 573

QY 301 GTCTCTGTGACTTCTTTTACCCCTATGCTGACCATATCCACAGCGTGGAGATGAAGCAGG 360

Db 574 GCCTCTCCAATTTCTTCGATTCCGTTGCCGACCGTATCCATTCGCTGGAGATGCAGAAG 633

QY 361 GCGCTACTTTGCTGCACTGTGCTGCTGTGAGCCGCTCAGCTGCGCTGCTGCTGCTGCTGCT 420

Db 634 GCGCACACTGTGTCATTTGCTGCTGGGTGAGCCGCTCAGCTGCGCTGCTGCTGCTGCTGCT 693

QY 421 ACCTCATGAAGTACACAGCCCATGTCCCTGTGGACGCCACACACGTGGACCAAGTCAATGCC 480

Db 694 ACCTCATGAAGTACCATGCCATGTCCCTGTAGATGCCACACTTGGACCAAGTCAATGCCA 753

QY 481 GGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTATCCACTATGATGTTCCAAT 540

Db 754 GGCCCATCATCCGCGCCCAACAGTGGCTTTTGGGAGCAGCTATCCACTATGATGTTCCAAT 813

QY 541 TGTTCGCAAGACACATGTCACATGTCAGTTCCCGCAGTGGGAATGATCCCTGACATCT 600

Db 814 TCTTTGCAAGATACATCAATCAGATGATGACTGCCCAATGGGACGATCCACAGCATCT 873

QY 601 ATGAGAAGGAGTCCGTTTGTATGATTCCTGATGATCCACTGTGAGCATCCCGACGAGCCCTGCAAT 658

Db 874 ACGAGAAGGAGACCCGTTTGTATGATGATCCACTGTGAAGCAGCCCGACCAAGTCAATCCAGT 931

RESULT 11

AL807241 186768 bp DNA linear HTG 17-AUG-2002

LOCUS Mus musculus chromosome 11 clone RP23-453E10, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.

DEFINITION AL807241

ACCESSION AL807241.4 GI:22416080

VERSION HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 186768)

REFERENCE Lovell, J.

AUTHORS Direct Submission

TITLE Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

JOURNAL On Aug 21, 2002 this sequence version replaced gi:21698134.

COMMENT humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

Center: Wellcome Trust Sanger Institute

Center project name: BM453E10

Center project name: BM453E10

Center project name: Project Information

Center project name: Summary Statistics

Center project name: Assembly program: XGAP4; version 4.5

Center project name: Chemistry: Dye-terminator; 100% of reads

Center project name: Consensus quality: 186346 bases at least Q40

Center project name: Consensus quality: 186456 bases at least Q30

QV 181 AGCTCATGCTGTCTAGCAACAGATCACCATTGGTTCATAATGTCTCAGTGGAAGGTACTGA 240

[illegible]

RESULT 14	
ALI133545	
LOCUS	ALI133545 linear PRI 10-NOV-2000
DEFINITION	Human DNA sequence from clone RP11-386K14 on chromosome Xp11.23-11.4. Contains ESTs, STSs, GSSs and cpg islands. Contains a gene for a novel protein similar to a dual specificity phosphatase and the 5' end of the UTX gene for the ubiquitously transcribed X chromosome tetratrinucleotide repeat protein, complete sequence.
ACCESSION	ALI133545
VERSION	ALI133545.10 GI:9368987
KEYWORDS	HTG; Cpg island; phosphatase.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 140717)
TITLE	Howden, P.
JOURNAL	Direct Submission
COMMENT	Submitted (11-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk On Jul 22, 2000 this sequence version replaced gi:9367888. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sv., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>

RP11-386N14 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-386N14. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP11-386N14 is at 140717 in this sequence. The true right end of clone RP6-105D16 is at 100 in this sequence.

FEATURES

source

```
1. 140717
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="X"
   /map="p11.23-11.4"
   /clone="RP11-386N14"
   /clone_lib="RPCI-11.2"
```

repeat_region

```
1. 200
   /note="AluJb repeat: matches 101. .293 of consensus"
```

repeat_region

```
326. .633
   /note="AluY repeat: matches 5. .310 of consensus"
```

repeat_region

```
717. .941
   /note="MER20 repeat: matches 3. .217 of consensus"
```

misc_feature

```
complement(1062..1294)
   /note="match: GSS: Em:AQ629580"
```

repeat_region

```
1062. .1256
   /note="AluSg/x repeat: matches 117. .310 of consensus"
```

misc_feature

```
complement(1062..1245)
   /note="match: GSS: Em:AQ215148"
```

misc_feature

```
complement(1062..1243)
   /note="match: GSS: Em:AQ005454"
```

misc_feature

```
1066. .1245
   /note="match: GSS: Em:AQ487216"
```

repeat_region

```
1294. .1404
   /note="MER3 repeat: matches 50. .165 of consensus"
```

repeat_region

```
1429. .1580
   /note="AluJb repeat: matches 134. .285 of consensus"
```

repeat_region

```
1581. .1884
   /note="AluJo repeat: matches 2. .303 of consensus"
```

repeat_region

```
1887. .1925
   /note="MER3 repeat: matches 159. .205 of consensus"
```

repeat_region

```
1900. .1940
   /note="MER3 repeat: matches 37. .77 of consensus"
```

repeat_region

```
1942. .2090
   /note="MIR repeat: matches 29. .172 of consensus"
```

repeat_region

```
2091. .2377
   /note="AluSx repeat: matches 1. .294 of consensus"
```

repeat_region

```
2378. .2437
   /note="MIR repeat: matches 172. .232 of consensus"
```

repeat_region

```
2644. .2958
   /note="AluSp repeat: matches 1. .313 of consensus"
```

repeat_region

```
3022. .3323
   /note="AluSx repeat: matches 1. .300 of consensus"
```

repeat_region

```
3496. .3581
   /note="L1Mc5 repeat: matches 7932. .7925 of consensus"
```

repeat_region

```
3698. .3743
   /note="23 copies 2 mer tt 73% conserved"
```

repeat_region

```
3749. .3969
   /note="AluJo repeat: matches 82. .296 of consensus"
```

repeat_region

```
4090. .4207
   /note="L1P repeat: matches 3626. .3747 of consensus"
```

repeat_region

```
4228. .4273
   /note="23 copies 2 mer tt 73% conserved"
```

repeat_region

```
4276. .4625
   /note="L1PA5 repeat: matches 5795. .6145 of consensus"
```

repeat_region

```
4626. .4687
   /note="L1PA5 repeat: matches 5744. .5805 of consensus"
```

repeat_region

```
4802. .5089
   /note="AluSp repeat: matches 1. .289 of consensus"
```

repeat_region

```
5528. .5820
   /note="AluSg repeat: matches 1. .293 of consensus"
```

repeat_region

```
6003. .6306
   /note="AluSg repeat: matches 1. .304 of consensus"
```

repeat_region

```
6325. .6618
   /note="AluJb repeat: matches 1. .298 of consensus"
```

repeat_region

```
6890. .7198
   /note="AluSx repeat: matches 1. .308 of consensus"
```

repeat_region

```
7303. .7611
   /note="AluSg repeat: matches 1. .309 of consensus"
```

repeat_region

```
7669. .7965
   /note="AluSx repeat: matches 1. .305 of consensus"
```

repeat_region

```
7984. .8110
   /note="FLAM_C repeat: matches 1. .127 of consensus"
```

repeat_region

```
8237. .8405
   /note="AluSx repeat: matches 124. .294 of consensus"
```

repeat_region

```
8406. .8714
   /note="AluJb repeat: matches 1. .308 of consensus"
```

repeat_region

```
9006. .9517
   /note="L1MA5A repeat: matches 5773. .6291 of consensus"
```

repeat_region

```
10250. .10558
   /note="AluSg repeat: matches 3. .305 of consensus"
```

repeat_region

```
10560. .10695
   /note="AluSx repeat: matches 1. .136 of consensus"
```

misc_feature

```
complement(10653..11141)
   /note="match: STS: Em:HSC32F11"
```

repeat_region

```
10906. .11100
   /note="L2 repeat: matches 2483. .2698 of consensus"
```

repeat_region

```
11355. .11661
   /note="AluSg repeat: matches 2. .308 of consensus"
```

repeat_region

```
12351. .12478
   /note="FLAM_C repeat: matches 5. .133 of consensus"
```

repeat_region

```
12682. .12747
   /note="MIR repeat: matches 156. .220 of consensus"
```

repeat_region

```
12959. .13227
   /note="AluJo repeat: matches 1. .269 of consensus"
```

repeat_region

```
13293. .13459
   /note="MLTIF repeat: matches 369. .534 of consensus"
```

repeat_region

```
13544. .13740
   /note="MLTIF repeat: matches 87. .284 of consensus"
```

misc_feature

```
complement(13746..14141)
   /note="match: GSS: Em:AQ383008"
```

repeat_region

```
13755. .14080
   /note="MER51A repeat: matches 50. .371 of consensus"
```

repeat_region

```
14141. .14179
   /note="MER51A repeat: matches 1. .39 of consensus"
```

repeat_region

```
14242. .14290
   /note="MLTIF repeat: matches 33. .81 of consensus"
```

repeat_region

```
14462. .14768
   /note="AluSg repeat: matches 1. .307 of consensus"
```

repeat_region

```
14890. .15196
   /note="AluSg repeat: matches 1. .302 of consensus"
```

repeat_region

```
15237. .15530
   /note="AluSg repeat: matches 3. .297 of consensus"
```

repeat_region

```
15570. .15709
   /note="MER3 repeat: matches 26. .173 of consensus"
```

repeat_region

```
15712. .16101
   /note="MER3 repeat: matches 2168. .2585 of consensus"
```

repeat_region

```
16485. .16539
   /note="MIR repeat: matches 106. .148 of consensus"
```

repeat_region

```
16540. .16837
```

```
repeat_region /note="AluSg repeat: matches 1. .298 of consensus"
16838..16876
repeat_region /note="MIR repeat: matches 64..106 of consensus"
17090..17409
repeat_region /note="AluSg repeat: matches 1. .312 of consensus"
17596..17673
repeat_region /note="L2 repeat: matches 2678..2749 of consensus"
17799..18112
repeat_region /note="MLT1A1 repeat: matches 1. .365 of consensus"
18186..18563
repeat_region /note="THE1C repeat: matches 1. .371 of consensus"
18564..20138
repeat_region /note="THE1C-internal repeat: matches 1. .1580 of consensus"
20139..20462
repeat_region /note="THE1C repeat: matches 1. .314 of consensus"
20473..20779
repeat_region /note="AluSg repeat: matches 1. .309 of consensus"
20780..20955
repeat_region /note="L1MB4 repeat: matches 6005..6185 of consensus"
20984..21286
repeat_region /note="AluSx repeat: matches 1. .305 of consensus"
21681..21940
repeat_region /note="L1MC5 repeat: matches 7645..7913 of consensus"
22012..22301
repeat_region /note="AluSx repeat: matches 1. .290 of consensus"
22669..22750
repeat_region /note="MER2 repeat: matches 258..335 of consensus"

Query Match 43.5%; Score 362.6; DB 9; Length 140717;
Best Local Similarity 70.9%; Pred. No. 7.5e-88;
Matches 564; Conservative 0; Mismatches 199; Indels 32; Gaps 5;

QY 48 GACACATGACCCAGCCGCTTGATGACAGACACCCCTCGTGCCTTCCCA-----GTTTCAG 101
DB 40466 GACCTCGGACGACGCGCTTGATGACAGACATCCGCGCTCTCTTTTCATCATCTCAGGGT 40525
QY 102 TTCGGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAAAGCCGTATATCAGCAAT 161
DB 40526 GTCCAGCAGCCCTCCATCTACAGCTTCTCCCAATACCAAGAGCTGTGTTCTCAGCAAT 40585
QY 162 GGTGTGCCGCGCAACAAGCTCATGCTGTGTAGCAACCAAGATCACCATTGGTCATCAAT 221
DB 40586 GGTGTGCCGCGCAACGAACTCCTTCTGTCCAGCAATCGCATCCGCGCATTTGTCAAT 40645
QY 222 GTCTCAGTGAGGTAGTGAACACCTTGTATGAGGATATCAGTACATGCGATCCTGTG 281
DB 40646 GCCTCGGTGGAAGTGGTCAACGATTTCTCGAGGGCATTCAGTACATAAAGGTCCTGT 40705
QY 282 GTCATCTCCCTTAACCTACGCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341
DB 40706 ACCGATCTGTGACTCGGCTCTACGACTTTTTCAGCCCATTTGCTGATCTTATCCAC 40765
QY 342 AGCGTGAGATGAAGCAGGCGGCTGCTGTGCTGCTGTGCTGTGAGCGGCTCA 401
DB 40766 ACCATCATATGAGCAGGCGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40825
QY 402 GCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
DB 40826 GCCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40885
QY 462 ACCTGGACCAAGTATCATCGCGCCCATCATCGGACCCCAACAGCGGCTTTTGGGAGCAGCTC 521
DB 40886 ACATGGACCAAGTATCGCGCCCATCATCGGACCCCAACAGCGGCTTTTGGGAGCAGCTC 40945
QY 522 ATCCATATGAGTTTCCAAATTTGTTGGCAAGAACACTGTGCAATGTTGAGTCCCGGAGT 581
DB 40946 ATCAATATGAGTTTCCAAATTTGTTGGCAAGAACACTGTGCAATGTTGAGTCCCGGAGT 41005
QY 582 GGATGATCCCTGATCATGAGAGGAGTCCGTTGATGATTCACCTGTGAGCCATC 641
DB 41006 GGTAAATCCCTGATCATGAGAGGAGTCCGTTGATGATTCACCTGTGAGCCATC 41065
```

```
QY 642 CCACGAGCCCTGCACTTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATC 701
DB 41066 CC-----GGCCAGCCCTGACATCTGCCATCATCTTGCACCAAGA-C 41107
QY 702 CAACCTTGAACATCTACTTTTGTGATACAGAAAAAAGAGATGATGCTTTTATGAGC 761
DB 41108 TGAACCTTGAACACTGACATTTTGTAGTAAG--AAACCGGATGGTGCCTTGTAAAGG 41165
QY 762 AAAAAAAGA-----GTTGCTGTAGCTTTTAACTTTATATATCCATTTTTCAGATTA 815
DB 41166 GCAAGAAAAAGGAGGGGTTGAGGTTTGAAGTAGTAGCCCTTACCTTAATAGAT 41225
QY 816 AACTAATTGTGAGAT 830
DB 41226 AATTATGAACAT 41240

RESULT 15
LOCUS AX040320 935 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 1 from Patent WO063393.
ACCESSION AX040320
VERSION AX040320.1 GI:11230164
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Luche,R.M. and Wei,B.
TITLE Dsp-8 dual-specificity map kinase phosphatase
JOURNAL Patent: WO 0063393-A 1 26-OCT-2000;
Ceptyr, Inc. (US)
FEATURES
source
location/Qualifiers
1..935
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 242 a 253 c 221 g 219 t
ORIGIN

Query Match 43.5%; Score 362.4; DB 6; Length 935;
Best Local Similarity 73.1%; Pred. No. 3.3e-88;
Matches 481; Conservative 0; Mismatches 171; Indels 6; Gaps 1;

QY 48 GACACATGACCCAGCCGCTTGATGACAGACACCCCTCGTGCCTTCCCA-----GTTTCAG 101
DB 156 GACCTCGGACGACGCGCTTGATGACAGACATCCGCGCTCTCTTTTCATCATCTCAGGGT 215
QY 102 TTCGGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAAAGCCGTATATCAGCAAT 161
DB 216 GTCCAGCAGCCCTCCATCTACAGCTTCTCCCAATACCAAGAGCTGTGTTCTCAGCAAT 275
QY 162 GGTGTGCCGCGCAACAAGCTCATGCTGTGTAGCAACCAAGATCACCATTGGTCATCAAT 221
DB 276 GGTGTGCCGCGCAACGAACTCCTTCTGTCCAGCAATCGCATCCGCGCATTTGTCAAT 335
QY 222 GTCTCAGTGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATGCGAGTACCTGTG 281
DB 336 GCCTCGGTGGAAGTGGTCAACGATTTCTCGAGGGCATTCAGTACATAAAGGTCCTTGT 395
QY 282 GCTGACTCCCTTAACCTACGCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341
DB 396 ACCGATCTGTGACTCGGCTCTACGACTTTTTCAGCCCATTTGCTGATCTTATCCAC 455
QY 342 AGCGTGAGATGAAGCAGGCGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
DB 456 ACCATCATATGAGCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 402 GCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
DB 516 GCCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 462 ACCTGGACCAAGTATCATCGCGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC 521
```

```

  || ||||| |||| |||| ||||| ||||| ||||| ||||| ||||| |||||
Db 576 ACATGGACCAAGTCGCGCCGCCCATCATCCGCCCAACAACGCTTTTGGGAACAGCTC 635
QY 522 ATCCACTATGAGTTCCTCAATTGTTGGCAAGAACACTGTGCACATGGTCACTTCCCCAGTG 581
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 ATCAATTACGAATTCAAGCTGTTTAATAACAACACCGTGGCGCATGATCAACTGCCCGGTA 695
QY 582 GGAATGATCCCTGACATCTATGAGAAGGAAAGTCCGTTTGATGATTCACACTGTGAGCCATC 641
  || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 696 GGTAACATCCCTGACATCTATGAGAGGACCTAGGTATGATATCAATGTAGCCATC 755
QY 642 CCACGAGCCCTGCAATGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGA 699
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 CCGGCCAGCCCTGACATCTGCCATCGATCTTGCACCAAGACTGAACCTTGAACACTGA 813
```

Search completed: April 21, 2003, 17:45:07
Job time : 2384 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:40:53 ; Search time 184 Seconds
(without alignments)
10207.427 Million cell updates/sec

Title: US-09-527-376-1
Perfect score: 834
Sequence: 1 ctttctctgtatttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	834.	100.0	834	21	AAA96619
2	805.6	96.6	1140	22	AAS35372
3	805.6	96.6	1140	22	AAS35019
4	805.6	96.6	1394	22	AAI58479
5	805.6	96.6	1407	22	AAH22963
6	804	96.4	7537	22	AAS36235
7	804	96.4	7537	22	AAK65643
8	793	95.1	1783	22	AAI60265
9	704.6	84.5	775	22	AAF63573

c	10	496	59.5	496	22	ABA48128
c	11	496	59.5	496	22	ABA66007
c	12	496	59.5	496	22	AAK14429
c	13	496	59.5	496	22	AAK40164
c	14	496	59.5	496	22	AAI20936
c	15	496	59.5	496	22	AAI46179
c	16	496	59.5	496	22	AAI06647
c	17	496	59.5	496	24	ABSI4218
c	18	463	55.5	463	22	ABA42999
c	19	463	55.5	463	22	ABA53415
c	20	463	55.5	463	22	AAK01684
c	21	463	55.5	463	22	AAK27133
c	22	463	55.5	463	22	AAI11722
c	23	463	55.5	463	22	AAI33028
c	24	463	55.5	463	22	AAI01650
c	25	463	55.5	463	22	ABSO1688
c	26	370.4	44.4	908	22	AAF63572
c	27	362.4	43.5	935	21	AAC67878
c	28	361	43.3	969	22	AAK52001
c	29	311.2	37.3	894	22	AAK52985
c	30	178.6	21.4	1466	20	AAH8831
c	31	177.6	21.3	1137	21	AAH88550
c	32	177.6	21.3	1729	20	AAZ28369
c	33	177.6	21.3	24000	21	AAH8551
c	34	171.4	20.6	1406	21	AAZ52479
c	35	152.6	18.3	456	22	AAK43861
c	36	152.6	18.3	456	24	ABSI48090
c	37	123.6	14.8	347	22	AAS34897
c	38	98.4	11.8	467	22	AAK32630
c	39	98.4	11.8	467	24	ABSO7421
c	40	88.8	10.6	1933	24	ABI99789
c	41	88	10.6	2000	24	ABK84633
c	42	88	10.6	2000	24	ABN97279
c	43	88	10.6	2000	24	ABL69410
c	44	88	10.6	2000	24	ABK35490
c	45	87	10.4	1209	24	ABK39796

ALIGNMENTS

RESULT 1

AAA96619
ID AAA96619 standard; cDNA; 834 BP.

XX AAA96619;

AC AAA96619;

XX 08-FEB-2001 (first entry)

DE CDNA encoding a human dual-specificity phosphatase 2 (DSP-2).

XX Human; mitogen-activated protein kinase; MAP kinase; cell survival;
KW dual-specificity phosphatase; DSP-2; cell proliferation; DSP-2;
KW cell differentiation; cancer; graft-versus-host disease; allergy;
KW autoimmune disease; metabolic disease; cell cycle abnormality; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

CDS 69..635

FT /*tag= a

FT /product= "dual-specificity phosphatase 2 (DSP-2)"

XX WO200056899-A1.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07589.

XX 24-MAR-1999; 990US-0125957.

XX 16-MAR-2000; 2000US-0527376.

XX (CEPT-) CEPTYR INC.

XX Luche RM, Wei B;
XX WPI: 2000-638264/61.
DR P-PSDB; AAB19008.
XX
XX Novel dual-specificity mitogen activated protein (MAP) kinase
PT phosphatase polypeptide useful in assays to identify agents that
PT modulate the protein activity, which are used to treat cancer and
PT autoimmune diseases -
XX
XX Claim 7; Fig 1: 51pp; English.
XX
XX The present sequence encodes a human dual-specificity mitogen-
CC activated protein (MAP) kinase phosphatase (DSP-2). The protein has
CC the ability to dephosphorylate an activated MAP-kinase. The DSP-2
CC protein is used for treatment of conditions associated with cell
CC proliferation, cell differentiation, and cell survival. The DSP-2
CC protein is useful for screening for modulators. Agents that modulate
CC DSP-2 activity are useful for treating cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth, abnormal cell proliferation and cell cycle abnormalities.
XX
XX Sequence 834 BP; 197 A; 226 C; 183 G; 228 T; 0 other;
SQ
Query Match 100.0%; Score 834; DB 21; Length 834;
Best Local Similarity 100.0%; Pred. No. 1e-237;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTTCCTGATATTTTGTCTTCAATCTTGTGTTTCGTCGACGTGACCACTGACCCA 60
DB 1 CTTTTCCTGATATTTTGTCTTCAATCTTGTGTTTCGTCGACGTGACCACTGACCCA 60
QY 61 CCGCCTTGATGACAGACCCCTGCTGCTCCAGTTCAGTTCGCGAGCCCTCAGTCA 120
DB 61 CCGCCTTGATGACAGACCCCTGCTGCTCCAGTTCAGTTCGCGAGCCCTCAGTCA 120
QY 121 CGCGCCTCTCGCAGATACCAAAAGCCCTGTATATCAGCAATGGTGTGCGCCGCCCAACA 180
DB 121 CGCGCCTCTCGCAGATACCAAAAGCCCTGTATATCAGCAATGGTGTGCGCCGCCCAACA 180
QY 181 AGCTCATGCTCTGACAGATACCAAGATCAGCATGGTGTATCTCAGTGGAGTAGTGA 240
DB 181 AGCTCATGCTCTGACAGATACCAAGATCAGCATGGTGTATCTCAGTGGAGTAGTGA 240
QY 241 ACACCTTGATGAGGATATCCAGTACATGACAGTACCTGTGCTGACCTTCACTTAC 300
DB 241 ACACCTTGATGAGGATATCCAGTACATGACAGTACCTGTGCTGACCTTCACTTAC 300
QY 301 GTCCTGTGACTTCTTTGACCTATTGCTGACCATATCCAGCGGTGGAGATGAAGCAGG 360
DB 301 GTCCTGTGACTTCTTTGACCTATTGCTGACCATATCCAGCGGTGGAGATGAAGCAGG 360
QY 361 GCGGTACTTTCGTCGACGTGCTGCTGTGTGAGCGGTGACGTGCGCTGCTGCGCT 420
DB 361 GCGGTACTTTCGTCGACGTGCTGCTGTGTGAGCGGTGACGTGCGCTGCTGCGCT 420
QY 421 ACCTCATGACTACACGCCATGCTCCCTGCTGGAGCCGCCACAGTGGACCAAGTCATGCC 480
DB 421 ACCTCATGACTACACGCCATGCTCCCTGCTGGAGCCGCCACAGTGGACCAAGTCATGCC 480
QY 481 GGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCAAT 540
DB 481 GGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCAAT 540
QY 541 TGTTTGGCAAGAACTGTGACATGGTTCAGTTCCTCCAGTGGGAATGATCCCTGACATCT 600
DB 541 TGTTTGGCAAGAACTGTGACATGGTTCAGTTCCTCCAGTGGGAATGATCCCTGACATCT 600
QY 601 ATGAGAAGGAAGTCGGTTTGTATGATTCACCTGTGAGCCATCCACGAGCCCTGCTATGG 660
DB 601 ATGAGAAGGAAGTCGGTTTGTATGATTCACCTGTGAGCCATCCACGAGCCCTGCTATGG 660

QY 661 AGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAAACTTGAACATCTTACT 720
DB 661 AGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAAACTTGAACATCTTACT 720
QY 721 TTTGTTGATACAGAAAAACAGATGATGCTTTTATGACACAAAAAGAGTTGCTGTA 780
DB 721 TTTGTTGATACAGAAAAACAGATGATGCTTTTATGACACAAAAAGAGTTGCTGTA 780
QY 781 GCCTTTAACTTTTATAATCCATTTTTCAGATTAAACTAATTTGAGATGGTG 834
DB 781 GCCTTTAACTTTTATAATCCATTTTTCAGATTAAACTAATTTGAGATGGTG 834

RESULT 2

AAS35372

ID AAS35372 standard; cDNA; 1140 BP.

XX AAS35372;

AC AAS35372;

XX 17-DEC-2001 (first entry)

DT 17-DEC-2001 (first entry)

XX Human cardiovascular system antigen cDNA polynucleotide SEQ ID No 257.

DE Human cardiovascular system antigen cDNA polynucleotide SEQ ID No 257.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;

KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;

KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;

KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;

KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;

KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;

KW cerebrovascular disorder; nervous system disorder; bacterial infection;

KW fungal infection; viral infection; ocular disorder; endocrine disorder;

KW gastrointestinal disorder; renal disorder; respiratory disorder;

KW wound healing; skin aging; organ transplantation; tissue regeneration;

KW anti-infertility.

XX Homo sapiens.

OS Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01340.

XX 31-JAN-2000; 2000US-0179055.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 13-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-451930/48.
DR P-PSDB; AAU22098.
XX
PT New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system -
XX
XX Claim 1; SEQ ID No 257; 674pp; English.
XX
CC Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the
CC cardiovascular system antigen polypeptides of the invention.
CC Cardiovascular system antigens and their associated polynucleotides are
CC useful in the diagnosis, treatment and prevention of various types of
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. A pathological condition can be determined by
CC detecting the presence or absence of a mutation in a cardiovascular
CC system antigen polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma and
CC pleurisy. The polypeptides can also be used to aid wound healing, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues and in chemotaxis.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 96.6%; Score 805.6; DB 22; Length 1140;
Best Local Similarity 99.5%; Pred. No. 3.4e-229;
Matches 808; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 696 CGGCTTTGGGAGCAGCTATCCACTATGAGTTCCAAATGTTGGCAAGAACACTGTGCA 755
Qy 563 CATGGTCAGTCCCGACGATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 622
Db 756 CATGGTCAGTCCCGACGATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 815
Qy 623 GATTCCTAGTGGAGCCATCCACGAGCCCTGCAATGAGAGTCAGAGTACAGATCTATTG 682
Db 816 GATTCCTAGTGGAGCCATCCACGAGCCCTGCAATGAGAGTCAGAGTACAGATCTATTG 875
Qy 683 TTGATCTTACCAAGATCCAACTTCAACATCTACTTTTGTGTACAGAAAAAACA 742
Db 876 TTGATCTTACCAAGATCCAACTTCAACATCTACTTTTGTGTACAGAAAAAACA 935
Qy 743 GATGATGCCCTTTTATGAGCAAAAAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 802
Db 936 GATGATGCCCTTTTATGAGCAAAAAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 995
Qy 803 TTTTTCAGATTAACTAAATGTGAGATGGTG 834
Db 996 TTTTTCAGATTAACTAAATGTGAGATGGTG 1027

RESULT 4
AAI58479
ID AAI58479 standard; cDNA; 1394 BP.
XX
AC AAI58479;
DT
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 682.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
PN
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0820312.
PR 03-AUG-2000; 2000US-0853450.
PR 14-SEP-2000; 2000US-0862191.
PR 19-OCT-2000; 2000US-08693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB: AAM39323.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
PS Claim 1; SEQ ID NO 682; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 1394 BP; 315 A; 385 C; 377 G; 317 T; 0 other;

Query Match 96.6%; Score 805.6; DB 22; Length 1394;
Best Local Similarity 99.5%; Pred. No. 3.9e-229;
Matches 808; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 23 CATCTTGGTGTTCGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCCCTC 82
Db 441 CAGCCTTCGTGCTGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCCCTC 500
Qy 83 GTGTGCTTCCCAAGTTTCAGTTCGGGAGCCCTCAGTCAGCGGGCCTCTCGCAGATAACCAA 142
Db 501 GTGTGCTTCCCAAGTTTCAGTTCGGGAGCCCTCAGTCAGCGGGCCTCTCGCAGATAACCAA 560
Qy 143 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACAAACAAAGCTCATGCTGTCTAGCAACCA 202
Db 561 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACAAACAAAGCTCATGCTGTCTAGCAACCA 620
Qy 203 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 262
Db 621 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 680
Qy 263 GTACATCAGAGTACCTGTGGCTGACTCCCTCACTCACTCACTGCTGTGACCTTCTTTGACCC 322
Db 681 GTACATCAGAGTACCTGTGGCTGACTCCCTCACTCACTCACTGCTGTGACCTTCTTTGACCC 740
Qy 323 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTTACTTGTGCTGCTGTC 382
Db 741 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTTACTTGTGCTGCTGTC 800
Qy 383 TGCTGTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
Db 801 TGCTGTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
Qy 443 GTCCTGCTGGAGCGCCACACAGTGGACCAAGTATGCGGCCCATCATCGAGCCCAACAG 502
Db 861 GTCCTGCTGGAGCGCCACACAGTGGACCAAGTATGCGGCCCATCATCGAGCCCAACAG 920
Qy 503 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 562
Db 921 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 980
Qy 563 CATGGTCAGTCCCGACGATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 622
Db 981 CATGGTCAGTCCCGACGATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 1040
Qy 623 GATTCCACTGTAGCCCATCCACGAGCCCTGCAATGAGTTCAGAGTCAGAGTACAGATCTATTG 682
Db 1041 GATTCCACTGTAGCCCATCCACGAGCCCTGCAATGAGTTCAGAGTCAGAGTACAGATCTATTG 1100
Qy 683 TTGATCTTACCAAGATCCAACTTGAACATCTACTTTTGTGTACAGAAAAAACA 742
Db 1101 TTGATCTTACCAAGATCCAACTTGAACATCTACTTTTGTGTACAGAAAAAACA 1160
Qy 743 GATGATGCCCTTTTATGAGCAAAAAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 802
Db 1161 GATGATGCCCTTTTATGAGCAAAAAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 1220

XX OS Homo sapiens.
XX PN WO200155321-A2.
XX XX
XX PD 02-AUG-2001.
XX PF
XX 17-JAN-2001; 2001WO-US01340.
XX 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180828.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 14-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 21-SEP-2000; 2000US-0234274.
XX PR 25-SEP-2000; 2000US-0234397.
XX PR 25-SEP-2000; 2000US-0234998.
XX 26-SEP-2000; 2000US-0235484.
XX PR 27-SEP-2000; 2000US-0235834.
XX PR 27-SEP-2000; 2000US-0235836.
XX PR 29-SEP-2000; 2000US-0236327.
XX PR 29-SEP-2000; 2000US-0236367.
XX PR 29-SEP-2000; 2000US-0236368.
XX PR 29-SEP-2000; 2000US-0236369.
XX PR 29-SEP-2000; 2000US-0236370.
XX PR 02-OCT-2000; 2000US-0236802.
XX PR 02-OCT-2000; 2000US-0237037.
XX PR 02-OCT-2000; 2000US-0237038.
XX PR 02-OCT-2000; 2000US-0237039.
XX PR 02-OCT-2000; 2000US-0237040.
XX PR 13-OCT-2000; 2000US-0239935.
XX PR 13-OCT-2000; 2000US-0239937.
XX PR 20-OCT-2000; 2000US-0240960.
XX PR 20-OCT-2000; 2000US-0241221.
XX PR 20-OCT-2000; 2000US-0241785.
XX PR 20-OCT-2000; 2000US-0241786.
XX PR 20-OCT-2000; 2000US-0241787.
XX PR 20-OCT-2000; 2000US-0241808.
XX PR 20-OCT-2000; 2000US-0241809.
XX PR 20-OCT-2000; 2000US-0241826.
XX PR 01-NOV-2000; 2000US-0244617.
XX PR 08-NOV-2000; 2000US-0246474.
XX PR 08-NOV-2000; 2000US-0246475.
XX PR 08-NOV-2000; 2000US-0246476.
XX PR 08-NOV-2000; 2000US-0246477.
XX PR 08-NOV-2000; 2000US-0246478.
XX PR 08-NOV-2000; 2000US-0246523.
XX PR 08-NOV-2000; 2000US-0246524.
XX PR 08-NOV-2000; 2000US-0246525.
XX PR 08-NOV-2000; 2000US-0246526.
XX PR 08-NOV-2000; 2000US-0246527.
XX PR 08-NOV-2000; 2000US-0246528.
XX PR 08-NOV-2000; 2000US-0246532.
XX PR 08-NOV-2000; 2000US-0246609.
XX PR 08-NOV-2000; 2000US-0246610.
XX PR 08-NOV-2000; 2000US-0246611.
XX PR 08-NOV-2000; 2000US-0246613.
XX PR 17-NOV-2000; 2000US-0249207.
XX PR 17-NOV-2000; 2000US-0249208.
XX PR 17-NOV-2000; 2000US-0249209.
XX PR 17-NOV-2000; 2000US-0249210.
XX PR 17-NOV-2000; 2000US-0249211.
XX PR 17-NOV-2000; 2000US-0249212.
XX PR 17-NOV-2000; 2000US-0249213.
XX PR 17-NOV-2000; 2000US-0249214.
XX PR 17-NOV-2000; 2000US-0249215.
XX PR 17-NOV-2000; 2000US-0249216.
XX PR 17-NOV-2000; 2000US-0249217.
XX PR 17-NOV-2000; 2000US-0249218.
XX PR 17-NOV-2000; 2000US-0249244.
XX PR 17-NOV-2000; 2000US-0249245.
XX PR 17-NOV-2000; 2000US-0249264.
XX PR 17-NOV-2000; 2000US-0249265.
XX PR 17-NOV-2000; 2000US-0249266.
XX PR 17-NOV-2000; 2000US-0249297.
XX PR 17-NOV-2000; 2000US-0249299.
XX PR 17-NOV-2000; 2000US-0249300.
XX PR 01-DEC-2000; 2000US-0250160.
XX PR 01-DEC-2000; 2000US-0250391.
XX PR 05-DEC-2000; 2000US-0251030.
XX PR 05-DEC-2000; 2000US-0251988.
XX PR 05-DEC-2000; 2000US-0256719.
XX PR 06-DEC-2000; 2000US-0251479.
XX PR 08-DEC-2000; 2000US-0251856.
XX PR 08-DEC-2000; 2000US-0251868.
XX PR 08-DEC-2000; 2000US-0251869.
XX PR 08-DEC-2000; 2000US-0251989.
XX PR 11-DEC-2000; 2000US-0251990.
XX PR 11-DEC-2000; 2000US-0254097.
XX PR 05-JAN-2001; 2001US-0259678.
XX XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM:

DR WPI; 2001-451930/48.

New cardiovascular system related polynucleotides and polypeptides,
useful for diagnosing, treating and/or preventing disorders of the
cardiovascular system -
PT

PS Claim 1; SEQ ID No 1735; 674pp; English.

Sequences AAS35741-1AAS36942 represent genomic DNA molecules, which encode the cardiovascular system antigen polypeptides of the invention. Cardiovascular system antigens and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or absence of a mutation in a cardiovascular system antigen polynucleotide. The treatable disorders include autoimmune diseases such as rheumatoid arthritis, hyperproliferative disorders such as neoplasms of the breast or liver, cardiovascular disorders such as cardiac arrest, cerebrovascular disorders such as cerebral ischaemia, nervous system disorders such as Alzheimer's disease, infections caused by bacteria, viruses and fungi, ocular disorders such as corneal infection, endocrine disorders such as premature labour and infertility, gastrointestinal disorders such as Crohn's disease, renal disorders such as glomerulonephritis and respiratory disorders such as asthma and pleurisy. The polypeptides can also be used to aid wound healing, to prevent skin aging due to sunburn, to maintain organs before transplantation, to regenerate tissues and in chemotaxis.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at pub.wipo.int/pub/published_pct_sequences.

Query Match	96.4%	Score 804;	DB 22;	Length 7537;
Best Local Similarity	99.4%	Pred. No. 3.1e-228;		
Matches 807; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

QY	23	CATTCTTGTTGTTTCGCTGACTGCTGACCATGACGCCACCGCCTTGATGACAGCACCCCTC	82
Db	3810		3869
		CAGCCTTCGTGTTGCGTGAAGCTGACCTTGACCCACCGCCTTGATGACAGCACCCCTC	
QY	83	GTGTGCTTTCCAGATTCACTTCCGCGAGCCCTCACTCAGCGGCCCTCTCGCAGATAAACAA	142
Db	3870		3929
		GTGTGCTTTCCAGTTCACTTCCGCGAGCCCTCACTCAGCGGCCCTCTCGCAGATAAACAA	
QY	143	AAGCCTGTATATCAGCAATGGTGTGGCGGCCAACAAACAGCTCATGCTGTAGCAACCA	202
Db	3930		3989
		AAGCCTGTATATCAGCAATGGTGTGGCGGCCAACAAACAGCTCATGCTGTAGCAACCA	
QY	203	GATCACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCCA	262
Db	3990		4049
		GATCACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCCA	
QY	263	GTACATGCAGGTACCTGTGGCTGACTCCCTTAATCACTCTCTGTGACTTCTTTTGACCC	322
Db	4050		4109
		GTACATGCAGGTACCTGTGGCTGACTCCCTTAATCACTCTCTGTGACTTCTTTTGACCC	
QY	323	TATTGCTGACCATATCCACAGCGTGGAGATGAACAGAGGCCGTACTTTGCTGCACCTGTC	382
Db	4110		4169
		TATTGCTGACCATATCCACAGCGTGGAGATGAACAGAGGCCGTACTTTGCTGCACCTGTC	
QY	383	TGCTGTGTGAGCGCTCAAGCTGCGCTGTGCGCTACCTCATGAAGTACCACGCCAT	442
Db	4170		4229
		TGCTGTGTGAGCGCTCAAGCTGCGCTGTGCGCTACCTCATGAAGTACCACGCCAT	
QY	443	GTCCCTGCTGGAGCCCCACAGCTGAGNACCAAGTCATGCCGGGCCCATCATCCAGCCACAG	502
Db	4230		4289
		GTCCCTGCTGGAGCCCCACAGCTGAGNACCAAGTCATGCCGGGCCCATCATCCAGCCACAG	
QY	503	CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCACAAATGTTTGGCAAGAACACTGTGCA	562

PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228824.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231444.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
PS Disclosure; SEQ ID NO 20455; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SQ Sequence 7537 BP; 1856 A; 1859 C; 1786 G; 2036 T; 0 other;

Query Match 96.4%; Score 804; DB 22; Length 7537;
Best Local Similarity 99.4%; Pred. No. 3.1e-228;
Matches 807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 23 CATCTTGTGTTGCTGACTGACCTGACCCACCCGCTTGATGACAGCACCCCTC 82

OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
XX	
PR	25-APR-2000; 2000US-0552317.
XX	
PR	09-JUL-2000; 2000US-0598042.
XX	
PR	19-JUL-2000; 2000US-0620312.
XX	
PR	03-AUG-2000; 2000US-0653450.
XX	
PR	14-SEP-2000; 2000US-0662191.
XX	
PR	19-OCT-2000; 2000US-0693036.
XX	
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX	
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
WI	PI; 2001-442253/47.
DR	P-ESDB; AAM41109.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
XX	
PS	such as central nervous system injuries -
XX	
PS	Claim 1; SEQ ID NO 4254; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX	
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
XX	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
XX	
CC	of the invention may be used to treat diseases of the peripheral nervous
XX	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
XX	
CC	localised neuropathies and central nervous system diseases, such as
XX	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX	
CC	utilisation of the activities such as: immune system suppression,
XX	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
XX	
CC	C.N.S disorders.
XX	
CC	Note: The sequence data for this patent did not form part of the printed
XX	
CC	specification.
XX	
SQ	Sequence 1783 BP; 405 A; 477 C; 457 G; 436 T; 8 other;
	Query Match 95.1%; Score 793; DB 22; Length 1783;
	Best Local Similarity 99.3%; Pred. No. 2.5e-225;
	Matches 807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY	23 CATTCCTGGTGTTCGCTGACTGCTGACACATGCACCACCGCCTTGATGACAGCACCCCTC 82
DB	441 CAGCCTTCGTGCTTCGCTGACTGCTGACCACTGACCACTGCCACCGCCTTGATGACAGCACCCCTC 500
QY	83 GTGTGCGCTTCCCAGTTTCAGTTCGGCAGCCCTCAGTCAGCGCGCCTCTCGCAGATAACCAA 142
DB	501 GTGTGCGCTTCCCAGTTTCAGTTCGGCAGCCCTCAGTCAGCGCGCCTCTCGCAGATAACCAA 560
QY	143 AAGCGTGATATCAGCAATGGTGTGCGCGCCAACAACAGCTCATGCTCTGTAGCAACCA 202
DB	561 AAGCGTGATATCAGCAATGGTGTGCGCGCCAACAACAGCTCATGCTCTGTAGCAACCA 620
QY	203 GATCACCATTGGTCATPAAATGTTCTAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 262
DB	621 GATCACCATTGGTCATPAAATGTTCTAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 680
QY	263 GTACATGCAGGTACTGTGGGTGACTCCCCCTAACCTACGCTCTCTGTGACTTCTTTGACCC 322
DB	681 GTACATGCAGGTACTGTGGGTGACTCCCCCTAACCTACGCTCTCTGTGACTTCTTTGACCC 740

AAI60265

XX

3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527

Human pol

Human: nootropic; immunosuppressant;

Accession	Gene	Accession	Gene
U00096	<i>Escherichia coli</i> K12	U00096	<i>Escherichia coli</i> K12
621	GATCACCATGGTTCATCAATGTTCTAGTGGAGGTAGTAACACCTGTTATGAGGATATCCA	681	GTACATGAGGATACCTGTGGCTGACTCCCTAACCTACAGTCTCTGTGACTTCTTTGACCC
Db		Db	
263	GTACATGAGGATACCTGTGGCTGACTCCCTAACCTACAGTCTCTGTGACTTCTTTGACCC	740	GTACATGAGGATACCTGTGGCTGACTCCCTAACCTACAGTCTCTGTGACTTCTTTGACCC
Qy		Qy	

QY 323 TATTGCTGACCATATCCACAGCGGTGAGATGAAGCAGGCGG-TACTTTGCTGCACTGTG 381
DB 741 TATTGCTGACCATATCCACAGCGGTGAGATGAAGCAGGCGGTACTTTGCTGCACTGTG 800
QY 382 CTGCTGTGTGAGCGGCTCAGCTGCGCTGTGCTGCGCTACCTCATGAAGTACCACGCCA 441
DB 801 CTGCTGTGTGAGCGGCTCAGCTGCGCTGTGCTGCGCTACCTCATGAAGTACCACGCCA 860
QY 442 TGTCCCTGCTGGAGCGCCACAGCTGGACCAAGTCAATCGCGGCCCATATCGGACCCCAACA 501
DB 861 TGTCCCTGCTGGAGCGCCACAGCTGGACCAAGTCAATCGCGGCCCATATCGGACCCCAACA 920
QY 502 GCGGCTTTTGGGAGCAGCTATCCACATGAGTTCCCAATTTTGGCAAGAACACTGTGC 561
DB 921 GCGGCTTTTGGGAGCAGCTATCCACATGAGTTCCCAATTTTGGCAAGAACACTGTGC 980
QY 562 ACATGGTCAAGTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGAAAGTCCGTTTGA 621
DB 981 ACATGGTCAAGTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGAAAGTCCGTTTGA 1040
QY 622 TGATTCACATGTGAGCCATCCAGAGCCCTGCATTTGGAGTCAAGAGTACAGATCTATT 681
DB 1041 TGATTCACATGTGAGCCATCCAGAGCCCTGCATTTGGAGTCAAGAGTACAGATCTATT 1100
QY 682 GTTGATCTTACACCAAGATCCAACTTTGAACATTTACTTTTGTGATACAGAAAAAACC 741
DB 1101 GTTGATCTTACACCAAGATCCAACTTTGAACATTTACTTTTGTGATACAGAAAAAACC 1160
QY 742 AGATGATGCTTTTATGAGCACAAGAGTGTCTGTAGCTTTTAACTTTTAAATCCAT 801
DB 1161 AGATGATGCTTTTATGAGCACAAGAGTGTCTGTAGCTTTTAACTTTTAAATCCAT 1220
QY 802 TTTTTCAGATTAACATTAATTTGAGATGGTG 834
DB 1221 TTTTTCAGATTAACATTAATTTGAGATGGTG 1253

RESULT 9

ID AAF63573 standard; cDNA; 775 BP.

XX AAF63573;

AC AAF63573;

DT 11-MAY-2001 (first entry)

XX Human phosphatase AA915932_h coding sequence.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma; ss.

XX Homo sapiens.

OS WO200112819-A2.

PN 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

PF 13-AUG-1999; 99US-0149005.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI; 2001-211226/21.

XX P-PSDB; AAB73221.

XX New protein phosphatase polypeptide for diagnosing and treating

PT phosphatase related disorders such as cancer, schizophrenia, cardiac

PT dysfunction and/or vascular disorders

XX Claim 1; Fig 4; 138pp; English.

CC The present invention relates to phosphatase proteins and coding
sequences. The present sequence is one such phosphatase coding sequence.
CC Phosphatases are enzymes that catalyse the phosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.

XX SQ Sequence 775 BP; 179 A; 229 C; 180 G; 187 T; 0 other;

Query Match 84.5%; Score 704.6; DB 22; Length 775;

Best Local Similarity 99.4%; Pred. No. 3.2e-199;

Matches 707; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CATTCCTTGTGTTTTCGCTGACTGCTGACCACTGACCCACCGCCCTTGATGACAGCACCCCTC 82

DB 65 CAGCCTTCGCTGCTGCTGACTGCTGACCACTGACCCACCGCCCTTGATGACAGCACCCCTC 124

QY 83 GTGTGCTTCCCGATTCAGTTCCGGCAGCCCTCAGTCAGCGGCCCTTCGCGAGATAACCAA 142

DB 125 GTGTGCTTCCCGATTCAGTTCCGGCAGCCCTCAGTCAGCGGCCCTTCGCGAGATAACCAA 184

QY 143 AAGCCTGTATATCAGCAATGTTGGCCGCCAACAAAGCTCATGCTGTACACCA 202

DB 185 AAGCCTGTATATCAGCAATGTTGGCCGCCAACAAAGCTCATGCTGTACACCA 244

QY 203 GATCACCATTGCTCATCAATGCTCAGTGGAGTAGTGAACACCTTGTATGAGGATATCCA 262

DB 245 GATCACCATTGCTCATCAATGCTCAGTGGAGTAGTGAACACCTTGTATGAGGATATCCA 304

QY 263 GTACATGAGGTACCTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 322

DB 305 GTACATGAGGTACCTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 364

QY 323 TATTGCTGACCATATCCACAGCGTGGAGATCAAGCAGCGCGTACTTTGCTGCACTGTGC 382

DB 365 TATTGCTGACCATATCCACAGCGTGGAGATCAAGCAGCGCGTACTTTGCTGCACTGTGC 424

QY 383 TGTGCTGTGAGCGGCTCAGCTGCCCTGTGCTGCGCTAGCTCATGAAGTACCACGCCAT 442

DB 425 TGTGCTGTGAGCGGCTCAGCTGCCCTGTGCTGCGCTAGCTCATGAAGTACCACGCCAT 484

QY 443 GTCCTCTGCTGGAGCCCAACAGTGGAGTCAAGTCAATGCGGCCCATCATCGACCCCAACAG 502

DB 485 GTCCTCTGCTGGAGCCCAACAGTGGAGTCAAGTCAATGCGGCCCATCATCGACCCCAACAG 544

QY 503 CGGCTTTTGGGAGCAGCTCATCCACTATGATTCCTCAATTTGTTGGCAAGAACTGTGCA 562

DB 545 CGGCTTTTGGGAGCAGCTCATCCACTATGATTCCTCAATTTGTTGGCAAGAACTGTGCA 604

QY 563 CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGAT 622

DB 605 CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGAT 664

QY 623 GATTCCACTGTGAGCCCATCCACAGCCCTGCTGATGAGTCAAGTACAGATCTATTG 682

DB 665 GATTCCACTGTGAGCCCATCCACAGCCCTGCTGATGAGTCAAGTACAGATCTATTG 724

QY 683 TTGATCTTACACCAAGATCCAACTTGAACATTTACTTTTGTGATACAG 733

DB 725 TTGATCTTACACCAAGATCCAACTTGAACATTTACTTTTGTGATACAG 775

CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

XX Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;

Query Match 59.5%; Score 496; DB 22; Length 496;
Best Local Similarity 100.0%; Pred. No. 3.le-137;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTCGGCAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCCTGTATATACGCAAT 161
DB 496 TTCGGCAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCCTGTATATACGCAAT 437
QY 162 GGTGTGGCCGCCAACAAAGCTCATGCTGTCTAGCAACAGATACCAATGGTCAAT 221
DB 436 GGTGTGGCCGCCAACAAAGCTCATGCTGTCTAGCAACAGATACCAATGGTCAAT 377
QY 222 GTCTCAGTGAGTAGTGAACACCTTGTATGAGGATATCCAGTACATGCAAGGTACCTGTG 281
DB 376 GTCTCAGTGAGTAGTGAACACCTTGTATGAGGATATCCAGTACATGCAAGGTACCTGTG 317
QY 282 GCTGACTCCCTCAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341
DB 316 GCTGACTCCCTCAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 257
QY 342 AGCGTGGAGATGAAGCAGGCGCCTACTTTGCTGCACCTGTGCTGCTGTGAGCGCCCTCA 401
DB 256 AGCGTGGAGATGAAGCAGGCGCCTACTTTGCTGCACCTGTGCTGCTGTGAGCGCCCTCA 197
QY 402 GCTGCGCCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 461
DB 196 GCTGCGCCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 137
QY 462 ACCTGGACCAAGTCAATGCGCGCCCATCATCCGACCCAAACAGCGGCTTTTGGGAGCAGCTC 521
DB 136 ACCTGGACCAAGTCAATGCGCGCCCATCATCCGACCCAAACAGCGGCTTTTGGGAGCAGCTC 77
QY 522 ATCCACTATGAGTTCCAATTTTGGCAAGAACACTGTGCACATGCTGTCAGTTCCTCCAGTG 581
DB 76 ATCCACTATGAGTTCCAATTTTGGCAAGAACACTGTGCACATGCTGTCAGTTCCTCCAGTG 17
QY 582 GGAATGATCCCTGACA 597
DB 16 GGAATGATCCCTGACA 1

RESULT 12

AAK14429/C

ID AAK14429 standard; DNA; 496 BP.

XX

AC AAK14429;

XX

DT 05-NOV-2001 (first entry)

XX

DE Human brain expressed single exon probe SEQ ID NO: 14420.

XX

KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200157275-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00667.

XX

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains
XX

PS Example 4; SEQ ID NO: 14420; 650pp + Sequence Listing; English.

XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

XX

XX Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;

Query Match 59.5%; Score 496; DB 22; Length 496;

Best Local Similarity 100.0%; Pred. No. 3.le-137;

Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTCGGCAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCCTGTATATACGCAAT 161
DB 496 TTCGGCAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCCTGTATATACGCAAT 437
QY 162 GGTGTGGCCGCCAACAAAGCTCATGCTGTCTAGCAACAGATACCAATGGTCAAT 221
DB 436 GGTGTGGCCGCCAACAAAGCTCATGCTGTCTAGCAACAGATACCAATGGTCAAT 377
QY 222 GTCTCAGTGAGTAGTGAACACCTTGTATGAGGATATCCAGTACATGCAAGGTACCTGTG 281
DB 376 GTCTCAGTGAGTAGTGAACACCTTGTATGAGGATATCCAGTACATGCAAGGTACCTGTG 317
QY 282 GCTGACTCCCTCAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341
DB 316 GCTGACTCCCTCAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 257
QY 342 AGCGTGGAGATGAAGCAGGCGCCTACTTTGCTGCACCTGTGCTGCTGTGAGCGCCCTCA 401
DB 256 AGCGTGGAGATGAAGCAGGCGCCTACTTTGCTGCACCTGTGCTGCTGTGAGCGCCCTCA 197
QY 402 GCTGCGCCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 461
DB 196 GCTGCGCCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 137
QY 462 ACCTGGACCAAGTCAATGCGCGCCCATCATCCGACCCAAACAGCGGCTTTTGGGAGCAGCTC 521
DB 136 ACCTGGACCAAGTCAATGCGCGCCCATCATCCGACCCAAACAGCGGCTTTTGGGAGCAGCTC 77
QY 522 ATCCACTATGAGTTCCAATTTTGGCAAGAACACTGTGCACATGCTGTCAGTTCCTCCAGTG 581
DB 76 ATCCACTATGAGTTCCAATTTTGGCAAGAACACTGTGCACATGCTGTCAGTTCCTCCAGTG 17
QY 582 GGAATGATCCCTGACA 597
DB 16 GGAATGATCCCTGACA 1

RESULT 13

AAK40164/C

ID AAK40164 standard; DNA; 496 BP.

XX

AAK40164;
06-NOV-2001 (first entry)
Human bone marrow expressed single exon probe SEQ ID NO: 14721.
Human; bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukaemia; lymphoma; myeloma; ss.
Homo sapiens.
WO200157276-A2.
09-AUG-2001.
30-JAN-2001; 2001WO-US00668.
04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-488900/53.
Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human bone marrow -
Example 4; SEQ ID NO: 14721; 658pp + Sequence Listing; English.
The present invention provides a number of single exon nucleic acid
probes which are derived from genomic sequences expressed in the human
bone marrow. They can be used to measure gene expression in bone marrow
samples, which may enable the improved diagnosis and treatment of cancers
such as lymphoma, leukemia and myeloma. The present sequence is one of
the probes of the invention.
Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;
Query Match 59.5%; Score 496; DB 22; Length 496;
Best Local Similarity 100.0%; Pred. No. 3.1e-137;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 102 TTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAGCCTGTATATCAGCAAT 161
Db 496 TTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAGCCTGTATATCAGCAAT 437
Qy 162 GGTGTGGCGCCACAAACAGCTCATGTCTAGCAACACAGATCACCATGGTCATCAAT 221
Db 436 GGTGTGGCGCCACAAACAGCTCATGTCTAGCAACACAGATCACCATGGTCATCAAT 377
Qy 222 GTCCTAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATCAGGTACCTGTG 281
Db 376 GTCCTAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATCAGGTACCTGTG 317
Qy 282 GCTGACTCCCTTACTACGCTCTGTGACTTCTTTGACCTATTGTCGACCATATCCAC 341
Db 316 GCTGACTCCCTTACTACGCTCTGTGACTTCTTTGACCTATTGTCGACCATATCCAC 257
Qy 342 AGCGTGGAGATGAACAGCGGCGCTACTTTGTGTGCTACTGTGTGCTGAGCGGCTCA 401
Db 256 AGCGTGGAGATGAACAGCGGCGCTACTTTGTGTGCTACTGTGTGCTGAGCGGCTCA 197
Qy 402 GCTGCCCTGTCCCTGCTTACTCATATGAGTACCAACGCGCATGTCCCTGCTGACGCCAC 461
Db 196 GCTGCCCTGTCCCTGCTTACTCATATGAGTACCAACGCGCATGTCCCTGCTGACGCCAC 137

Qy 462 ACGTGGACCAAGTCATCGCGGCCCATCATCGACCAACAGCGGCTTTTGGAGCAGCTC 521
Db 136 ACGTGGACCAAGTCATCGCGGCCCATCATCGACCAACAGCGGCTTTTGGAGCAGCTC 77
Qy 522 ATCCACTATGATGTTCCCAATTTTGGCAAGAACACTGTGCACATGGTCAGTTCGCCAGTG 581
Db 76 ATCCACTATGATGTTCCCAATTTTGGCAAGAACACTGTGCACATGGTCAGTTCGCCAGTG 17
Qy 582 GGAATGATCCCTGACA 597
Db 16 GGAATGATCCCTGACA 1
RESULT 14
AAI20936/C
ID AAI20936 standard; DNA; 496 BP.
XX AC AAI20936;
XX DT 12-OCT-2001 (first entry)
XX DE Probe #10869 for gene expression analysis in human cervical cell sample.
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human cervical epithelial cells -
XX PS Claim 25; SEQ ID NO 10869; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SNP). The present sequence is one such probe. The SNPs are derived
XX CC from human HeLa cells. The SNPs can be used to produce a single exon
XX CC microarray, which can be used for measuring human gene expression in a
XX CC sample derived from human cervical epithelial cells. By measuring gene
XX CC expression, the probes are therefore useful in grading and/or staging
XX CC of diseases of the cervix, notably cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;
Query Match 59.5%; Score 496; DB 22; Length 496;
Best Local Similarity 100.0%; Pred. No. 3.1e-137;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 102 TTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAGCCTGTATATCAGCAAT 161
Db 496 TTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAAGCCTGTATATCAGCAAT 437

```
QY 162 GGTGTGGCGCCCAACAAAGCTCATGCTCTAGCAACAGATCACCATTGTCATCAAT 221
    |||
Db 436 GGTGTGGCGCCCAACAAAGCTCATGCTCTAGCAACAGATCACCATTGTCATCAAT 377
    |||
QY 222 GTCTCAGTGGAGGTAGTGAACACCTTTGTATGAGGATATCCAGTACATGCGAGGTACCTGTG 281
    |||
Db 376 GTCTCAGTGGAGGTAGTGAACACCTTTGTATGAGGATATCCAGTACATGCGAGGTACCTGTG 317
    |||
QY 282 GCTGACTCCCTTAACATCAGCTCTCTGTGACTTCTTTGACCTATTGACCTATTGTCACCATATCCAC 341
    |||
Db 316 GCTGACTCCCTTAACATCAGCTCTCTGTGACTTCTTTGACCTATTGTCACCATATCCAC 257
    |||
QY 342 AGGTGGAGATGAGGAGGCGCTGACTTTGCTGCACCTGTGCTGCTGTGAGGCGCTCA 401
    |||
Db 256 AGGTGGAGATGAGGAGGCGCTGACTTTGCTGCACCTGTGCTGCTGTGAGGCGCTCA 197
    |||
QY 402 GCTGCCCTGTGCTCGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 461
    |||
Db 196 GCTGCCCTGTGCTCGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 137
    |||
QY 462 ACGTGGACCAAGTCAATGCGGCGCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC 521
    |||
Db 136 ACGTGGACCAAGTCAATGCGGCGCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC 77
    |||
QY 522 ATCCACTATGAGTTCCTCAATTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 581
    |||
Db 76 ATCCACTATGAGTTCCTCAATTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 17
    |||
QY 582 GGAATGATCCCTGACA 597
    |||
Db 16 GGAATGATCCCTGACA 1
    |||
```

RESULT 15

AAI46179/c

ID AAI46179 standard; DNA; 496 BP.

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

XX AC AAI46179;

```
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;
```

Query Match 59.5%; Score 496; DB 22; Length 496;

Best Local Similarity 100.0%; Pred. No. 3.1e-137;

Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTCGGCAGCCCTCAGTCAAGCGGCTCTCGCAGATAACCAAAAGCGCTGTATATCAGCAAT 161

Db 496 TTCGGCAGCCCTCAGTCAAGCGGCTCTCGCAGATAACCAAAAGCGCTGTATATCAGCAAT 437

QY 162 GGTGTGGCGCCCAACAAAGCTCATGCTCTAGCAACAGATCACCATTGTCATCAAT 221

Db 436 GGTGTGGCGCCCAACAAAGCTCATGCTCTAGCAACAGATCACCATTGTCATCAAT 377

QY 222 GTCTCAGTGGAGGTAGTGAACACCTTTGTATGAGGATATCCAGTACATGCGAGGTACCTGTG 281

Db 376 GTCTCAGTGGAGGTAGTGAACACCTTTGTATGAGGATATCCAGTACATGCGAGGTACCTGTG 317

QY 282 GCTGACTCCCTTAACATCAGCTCTCTGTGACTTCTTTGACCTATTGTCACCATATCCAC 341

Db 316 GCTGACTCCCTTAACATCAGCTCTCTGTGACTTCTTTGACCTATTGTCACCATATCCAC 257

QY 342 AGGTGGAGATGAGGAGGCGCTGACTTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 401

Db 256 AGGTGGAGATGAGGAGGCGCTGACTTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 197

QY 402 GCTGCCCTGTGCTCGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 461

Db 196 GCTGCCCTGTGCTCGCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 137

QY 462 ACGTGGACCAAGTCAATGCGGCGCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC 521

Db 136 ACGTGGACCAAGTCAATGCGGCGCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC 77

QY 522 ATCCACTATGAGTTCCTCAATTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 581

Db 76 ATCCACTATGAGTTCCTCAATTTGTCGCAAGACACACTGTGCACATGCTCAGTTCCTCCAGTG 17

QY 582 GGAATGATCCCTGACA 597

Db 16 GGAATGATCCCTGACA 1

Search completed: April 21, 2003, 17:05:22

Job time : 200 secs

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488897/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human placenta -

Claim 25; SEQ ID No 14865; 654pp; English.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 17:45:14 ; Search time 88 Seconds
(without alignments)
9520.508 Million cell updates/sec

Title: US-09-527-376-1
Perfect score: 834
Sequence: 1 ctttctgtgttttttgc.....aaactaatgtgagatggtg 834

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues
Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	805.6	96.6	1140	9	US-10-091-504-257
2	805.6	96.6	1140	9	US-09-764-869-257
3	804	96.4	7537	9	US-10-091-504-1735
4	804	96.4	7537	10	US-09-764-869-1735
5	496	59.5	496	10	US-09-864-761-26715
6	463	55.5	463	10	US-09-864-761-10073
7	361	43.3	923	12	US-10-044-205A-41
8	340.8	40.9	570	12	US-10-044-205A-43
9	152.6	18.3	456	10	US-09-864-761-30844
10	149.6	17.9	239	10	US-09-783-590-8474
11	98.4	11.8	467	10	US-09-864-761-14459
12	88	10.6	2000	10	US-09-969-708-276
13	88	10.6	2000	10	US-09-919-497-10
14	88	10.6	2000	10	US-09-880-107-3774
15	87	10.4	1209	9	US-09-902-941-1918
16	87	10.4	1209	9	US-09-849-626-1918
17	87	10.4	1209	9	US-10-017-754-1918
18	87	10.4	1238	9	US-09-736-457-803
19	87	10.4	1238	9	US-09-902-941-803

ALIGNMENTS

RESULT 1
US-10-091-504-257
; Sequence 257, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 257
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1104)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-504-257

Query Match	96.6%	Score	805.6	DB	9	Length	1140
Best Local Similarity	99.5%	Pred. No.	1.1e-244				
Matches	808	Conservative	0	Mismatches	4	Indels	0
Gaps	0						
QY	23	CATTCTTGTGTTTCGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCCCTC	82				
Db	216	CAGCCTTCGTGCTGCTGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCCCTC	275				
QY	83	GTGTGCTTCCAGCTTCAGTTCGGGAGCGGCTTCGAGGCGCTTCGAGATACCA	142				
Db	276	GTGTGCTTCCAGCTTCAGTTCGGGAGCGGCTTCGAGGCGCTTCGAGATACCA	335				
QY	143	AAGCTGTATATACAGCAATGGTGGCCGCGCAACAACAGCTCATGCTGTGTAGCAACCA	202				
Db	336	AAGCTGTATATACAGCAATGGTGGCCGCGCAACAACAGCTCATGCTGTGTAGCAACCA	395				
QY	203	GATCACCATTGGTCATCAATGTCTCTAGTGAGGTAGTGAACACCTTGATGAGGATATCCA	262				
Db	396	GATCACCATTGGTCATCAATGTCTCTAGTGAGGTAGTGAACACCTTGATGAGGATATCCA	455				

Sequence 803, App
Sequence 803, App
Sequence 825, App
Sequence 825, App
Sequence 825, App
Sequence 825, App
Sequence 826, App
Sequence 826, App
Sequence 826, App
Sequence 826, App
Sequence 29, Appl
Sequence 801, App
Sequence 801, App
Sequence 801, App
Sequence 801, App
Sequence 804, App
Sequence 804, App
Sequence 804, App
Sequence 804, App
Sequence 51, Appl
Sequence 204, App
Sequence 311, App
Sequence 311, App
Sequence 311, App
Sequence 311, App
Sequence 485, App


```

; SEQ ID NO 1735
; LENGTH: 7537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1735

Query Match          96.4%; Score 804; DB 10; Length 7537;
Best Local Similarity 99.4%; Pred. No. 1.4e-243;
Matches 807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 23 CATTCCTGGTGTTCGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCTC 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3810 CAGCCTTCGGTTCGCTGACTGCTGACCACTGACCCACCGCCTTGATGACAGCACCTC 3869

QY 83 GTGTGCCTTCGCCAGTTCAGTTCGGCGACGCCCTCAGTCAGGGGCCCTCTCGCAGATAACCAA 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3870 GTGTGCCTTCGCCAGTTCAGTTCGGCGACGCCCTCAGTCAGGGGCCCTCTCGCAGATAACCAA 3929

QY 143 AAGCCTCTATATACGAATGGTGTGGCGCGCAACAAACAGCTCATGTGTCTAGCAACCA 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3930 AAGCCTCTATATACGAATGGTGTGGCGCGCAACAAACAGCTCATGTGTCTAGCAACCA 3989

QY 203 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3990 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 4049

QY 263 GTACATCGAGGTACTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4050 GTACATCGAGGTACTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 4109

QY 323 TATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCGGTACTTTGCTGCACTGTGC 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4110 TATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCGGTACTTTGCTGCACTGTGC 4169

QY 383 TGCTGGTGTGAGCGCTCAGCTGCGCTGTGCTCGCCTCACTCATGATGATACCAAGCCAT 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4170 TGCTGGTGTGAGCGCTCAGCTGCGCTGTGCTCGCCTCACTCATGATGATACCAAGCCAT 4239

QY 443 GTCCCTGCTGGACGCCCCACAGTGGACCAAGTCAATGCGGCCCATCATCCGACCCCAACAG 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4230 GTCCCTGCTGGACGCCCCACAGTGGACCAAGTCAATGCGGCCCATCATCCGACCCCAACAG 4289

QY 503 CGGCTTTTGGAGCAGGTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 562
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4290 CGGCTTTTGGAGCAGGTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 4349

QY 563 CATGGTCAGTTCCTCAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCCTTTGAT 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4350 CATGGTCAGTTCCTCAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCCTTTGAT 4409

QY 623 GATTCCACTGTGAGCCATCCACAGGCCCTTGCATTGGAGTCAGAGGTACAGATCTATTG 682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4410 GATTCCACTGTGAGCCATCCACAGGCCCTTGCATTGGAGTCAGAGGTACAGATCTATTG 4469

QY 683 TTGATCTTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAACA 742
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4470 TTGATCTTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAACA 4529

```

```

Db      4530 GATGATGCCCTTTTATGAGCACAAAAAAGAGTGTGCTTAGCTTTTAACTTTATAATCCATT 4589
            |||||
Qy      803 TTTTTTCAGATTAAACTAAATTGTCAGATGGTG 834
            |||||||
Db      4590 TTTTTTAAGATTAAACTAAATTGTCAGATGGTG 4621
            |||||

RESULT 5
US-09-864-761-26715/C
; Sequence 26715, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.

```

```

: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/006666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/006670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 26715
: LENGTH: 496
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL079299.8
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
: OTHER INFORMATION: EST_HUMAN HIT: BE222374.1, EVALUE 0.00e+00
: OTHER INFORMATION: SWISSPROT HIT: Q64623, EVALUE 2.00e-21
: OTHER INFORMATION: NT HIT: g11545790, EVALUE 3.00e-38
: IS-09-864-761-26715

```

Qy	162	GGTGTGGCCGCCAACAAACAAAGCTCATGCTGCTCTAGAACACAGATACACCATGGTTCATCAAT	221
Db	436	GGTGTGGCCGCCAACAAACAAAGCTCATGCTGCTCTAGCAACAGATACACCATGGTTCATCAAT	377
Qy	222	GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATGAGAGTACCTGTG	281
Db	376	GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATGAGAGTACCTGTG	317
Qy	282	GCTGACTCCCTAACTCACGTCCTGTGACTTCTTTGACCCTATTGCTGACCATATCCAC	341
Db	316	GCTGACTCCCTAACTCACGTCCTGTGACTTCTTTGACCCTATTGCTGACCATATCCAC	257
Qy	342	AGCGTGGAGATGAAGCAGGCGCGTACTTTGCTGCATGTGCTGCTGGTGTGAGCCGCTCA	401
Db	256	AGCGTGGAGATGAAGCAGGCGCGTACTTTGCTGCATGTGCTGCTGGTGTGAGCCGCTCA	197
Qy	402	GCTGCCCTGTGCGCTCGCGCTACCTCATGAAGTACACAGCCATGTCCTGTGTGAGACGCCAC	461
Db	196	GCTGCCCTGTGCGCTCGCGCTACCTCATGAAGTACACAGCCATGTCCTGTGTGAGACGCCAC	137
Qy	462	ACGTGGACCAAGTCATGCCGGCGCATCATCCGACCCACACAGCGGCTTTTGGGAGCAGCTC	521
Db	136	ACGTGGACCAAGTCATGCCGGCGCATCATCCGACCCACACAGCGGCTTTTGGGAGCAGCTC	77
Qy	522	ATCCACATATGAGTTCCTCAATTTTGGCAAGAACACATGTGCACATGTGTCTCCCAAGT	581
Db	76	ATCCACATATGAGTTCCTCAATTTTGGCAAGAACACATGTGCACATGTGTCTCCCAAGT	17
Qy	582	GGAATGATCCCTGACA	597
Db	16	GGAATGATCCCTGACA	1

RESULT. 6

```

US-09-864-761-10073/c
; Sequence 10073, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

```


Db 843 GCAAGAAAAAGGAGGGGTTGGAGTTTGAACGTAGTAAAGCCTTACCTTAATAGAATT 902
QY 816 AACTAAATTGTGAGAT 830
Db 903 AAATTCATGAACAT 917

RESULT 8

US-10-044-205A-43
; Sequence 43, Application US/10044205A
; Patent No. US20020123464A1
; GENERAL INFORMATION:
; APPLICANT: KAPELLER-LIBERMAN, Rosana
; APPLICANT: BANDARU, Rajasekhar
; TITLE OF INVENTION: 69087, 15821, and 15418, Methods and Compositions of Human Proteins
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 10147-5201
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 10/044,205A
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/242,428
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/241,884
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,877
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-044-205A-43

Query Match 40.9%; Score 340.8; DB 12; Length 570;
Best Local Similarity 75.5%; Pred. No. 1.5e-97;
Matches 423; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 73 CAGCACCTCGTGTGCTTCCCGTTCAGTTCCGCGCAGCCCTCAGTCAGCGCCCTCTCGC 132
Db 11 CGCGCTCTCTCTTTCATCATCTCAGGTGTCCAGCGCCCTCCATCTACAGCTTCTCC 70
QY 133 AGATAACCAAAAGCCTGTATACGAATGTGTGGCGCCCAACAAAGCTCATGCTGT 192
Db 71 AAATAACCAAGAGCTGTCTTCCAGCATGTGTGGCGCCCAACAAAGCTTCTGT 130
QY 193 CTAGCAACAGATACCATGTTGTCATCAATGTCTAGTGGAGGTAGTGAACACCTTTGATG 252
Db 131 CCAGCAATCGCATCACCGCCATTTGTAATGCTCGTGGTGAAGTGTCAACGTTATCTTCG 190
QY 253 AGGATATCCAGTACATGAGGTACCTGTGGTGTGACTCCCTTAACTCAGTCTCTGTGACT 312
Db 191 AGGCAATTCAGTACATAAGGTGCTGTACCGATGCTGTCGTCGCTCTACGACT 250
QY 313 TCTTTGACCCATTCTCTGCTGACCATATCCACAGGTGGAGTGAAGCAGGCGGTACTTTGC 372
Db 251 TTTTGGACCCATTCTCTGCTATCCACCATCATGATGAGGAGGCGGTACGCTGC 310
QY 373 TGCATGTGCTGTGGTGTGAGCGCCTCAGTGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
Db 311 TGCATGTGCTGTGGTGTGAGCGCCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
QY 433 ACCAGCCATGCTGCTGCTGAGCGCCACAGTGCAGCAAGTCAATGTCGCGCCCATCC 492
Db 371 ACCATCTCATGTGCTGCTGAGCGCCCATATGATGAGCAAGTGCAGCGCCCATCATCC 430
QY 493 GACCAACAGCGGCTTTTGGAGCAGCTCATCCATGATGATGATGATGATGATGATGATGATGATGAT 552
Db 431 GSCCAACAAACGCTTTTGGGAACAGCTCATCAATGATGATGATGATGATGATGATGATGATGATGAT 490
QY 553 ACAGTGTGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
Db 491 ACAGCTGCGCATGATCAACTCGCGGTAGGTAAACATCCCTGACATCTATGAGAGGACC 550

QY 613 TCCGTTTGATGATTCCTACTG 632
Db 551 TACGTACGATGATCAATG 570

RESULT 9

US-09-864-761-30844
; Sequence 30844, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30844
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004099.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
; OTHER INFORMATION: SWISSPROT HIT: O64623, EVALUE 3.00e-26
; OTHER INFORMATION: EST HUMAN HIT: BF130916.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: g111428005, EVALUE 0.00e+00
US-09-864-761-30844

Query Match 18.3%; Score 152.6; DB 10; Length 456;

[illegible]

```

RESULT 10
US-09-783-590-8474
; Sequence 8474, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12495
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8474
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (72)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (131)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

; NAME/KEY: misc feature
; LOCATION: (191)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-783-590-8474

Query Match          17.9%; Score 149.6; DB 10; Length 239;
Best Local Similarity 91.3%; Pred. No. 3.2e-37;
Matches 190; Conservative 0; Mismatches 13; Indels 5; Gaps

Qy 617 TTTGATGATCCACTGTGAGCCATCCACGAGGCCCTGCATTTGGAGTCAGAGGTACAGAT 676
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 TTTAATGATTCACGTGAAGCCATCCACGAGGCCCTGCANTGGAGTCAGAGGTACAGAT 91
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 677 CTATTTGTTGATCTTTACACCAAGATCCAAACTTGAACATTTCTACTTTGTTGATACAGAA 736
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 CTATTTGTTGATCTTTACACCAAGATCCAAACTTGAACATTTACTTTATTGGTACAGAA 751
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 737 AAAACAGATGATGCTTTT-ATGAGCACAAAAAGATT-GCTGTAGCTTTT---AACTT 791
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 AAAACAGATGATGCTTTTAAATGAGCACAAAAAAGNTTNGCTGTAGCTTGTGTAACITT 211
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 792 TATATCCATTTTTTTTTTCAGATTAAACT 819
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 AATATGCANTTTTTTTTTTCAGTTAAACT 239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11
US-09-864-761-14459
: Sequence 14459, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEI
: FILE REFERENCE: Aomic-a-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30

THIS PAGE BLANK (USPTO)